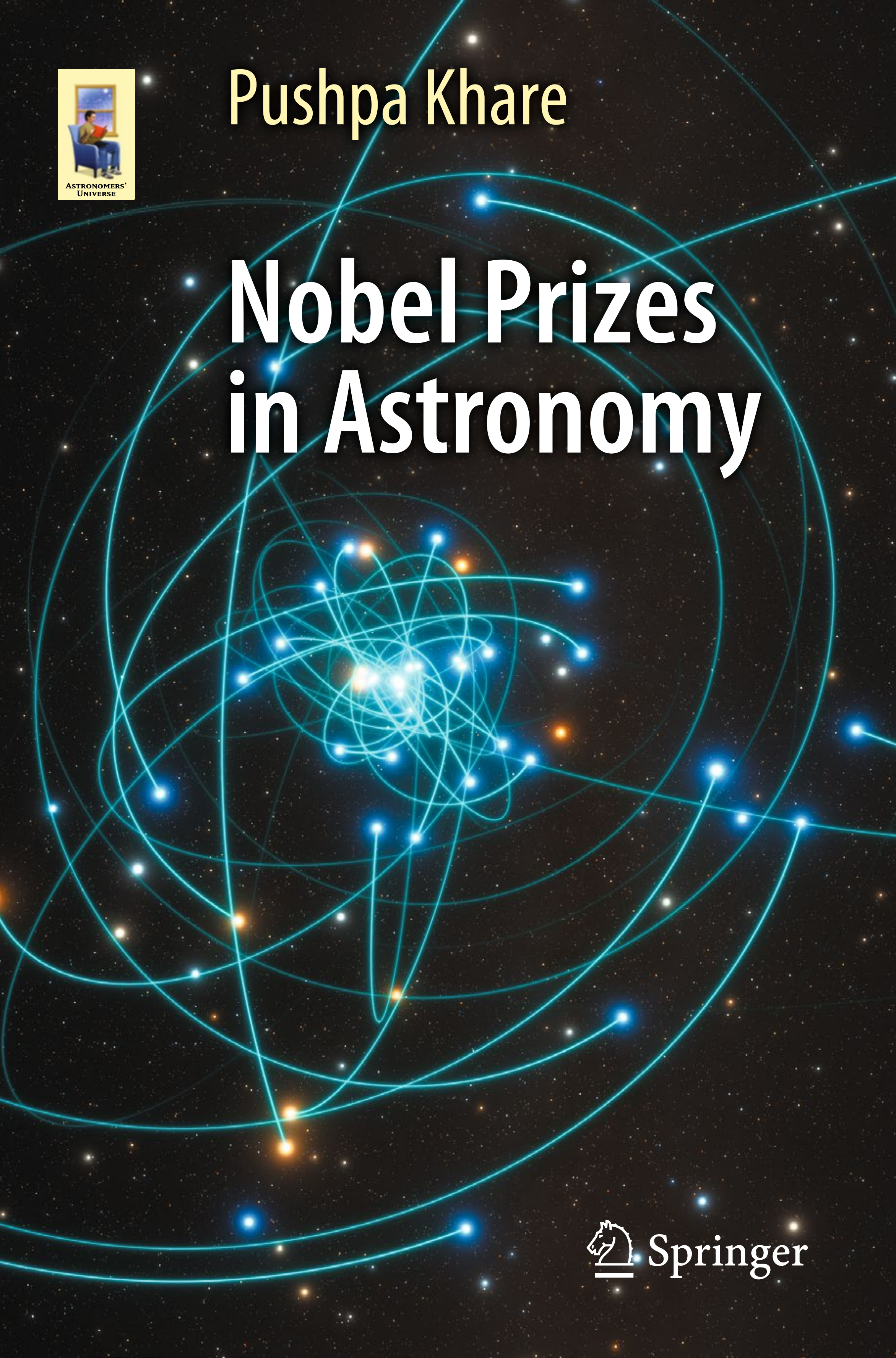


Astronomers' Universe

Pushpa Khare

Nobel Prizes in Astronomy

Springer

Astronomers' Universe

Series Editor
Martin Beech, Campion College, The University of Regina,
Regina, SK, Canada

The Astronomers' Universe series attracts scientifically curious readers with a passion for astronomy and its related fields. In this series, you will venture beyond the basics to gain a deeper understanding of the cosmos—all from the comfort of your chair.

Our books cover any and all topics related to the scientific study of the Universe and our place in it, exploring discoveries and theories in areas ranging from cosmology and astrophysics to planetary science and astrobiology.

This series bridges the gap between very basic popular science books and higher-level textbooks, providing rigorous, yet digestible forays for the intrepid lay reader. It goes beyond a beginner's level, introducing you to more complex concepts that will expand your knowledge of the cosmos. The books are written in a didactic and descriptive style, including basic mathematics where necessary.

Pushpa Khare

Nobel Prizes in Astronomy

Springer

Pushpa Khare
Pune, India

ISSN 1614-659X ISSN 2197-6651 (electronic)
Astronomers' Universe
ISBN 978-3-031-29638-3 ISBN 978-3-031-29639-0 (eBook)
https://doi.org/10.1007/978-3-031-29639-0

Cover illustration: This simulation shows the orbits of stars very close to the supermassive black hole at the heart of the Milky Way. One of these stars, named S2, orbits every 16 years and is passing very close to the black hole in May 2018. This is a perfect laboratory to test gravitational physics and specifically Einstein's general theory of relativity. Research into S2's orbit was presented in a paper entitled "Detection of the Gravitational Redshift in the Orbit of the Star S2 near the Galactic Centre Massive Black Hole", by the GRAVITY Collaboration, which appeared in the journal Astronomy & Astrophysics on 26 July 2018. Credit: ESO/L. Calçada/spaceengine.org

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

Astronomy is a subject which fascinates the young and old alike. Their interest is fed from time to time by press reports of some exciting astronomical events, often accompanied by evermore beautiful images made possible by continuously improving technology. I have witnessed the curiosity of high school students every time I gave a semi-popular hour long lecture to them, as the subsequent question-answer sessions invariably had to be forcibly stopped for want of time. So when my daughter Anupama suggested I write a book explaining the Nobel Prize winning works to students, I immediately took it up. I am very grateful to her for the suggestion. I have thoroughly enjoyed the process of writing. The material in the book is confined to Nobel Prizes in Astronomy, and it is targeted at readers who have a background of high school physics and are curious to learn about the research done at the forefront of astronomy. I have tried my best to write at a level understandable to them. Interested readers can always find more advanced material on the internet.

I am grateful to Prof. Ajit Kembhavi of IUCAA, Pune, for going through an initial draft of the manuscript and suggesting improvements which definitely helped. I am thankful to Prof. Rajaram Nityananda, Bangalore, for giving comments on some of the chapters and suggesting a few crucial changes. I have to thank my husband Avinash for his constant support in all my endeavours. His encouragement has been crucial for their successful completion. My son Apoorva went through the book patiently and corrected grammatical mistakes I had not thought I was capable of making. The figures, wherever not credited

in the text, have been made by Shashank Tarphe of IUCAA and Kaushal Sharma. I thank them for their patience and for making the figure corrections that I requested again and again. The portraits of two of the awardees have not been included in the book as the necessary permissions could not be obtained.

Pune, India Pushpa Khare
20 January 2023

Contents

1

Introduction

Recognition by others is one of the basic needs of human beings, and the recognition that comes from one's peers is definitely very satisfying. Scientists, artists, writers, social workers etc., are no exception to this. A large number of prizes have been instituted and periodically, often annually, awarded to honour and encourage the worthy among these. The biggest and best recognition that researchers in the fields of physics, chemistry, medicine, economics, social workers/activists/politicians and authors of literature contributing towards achieving peace in the world, can get is undoubtedly the Nobel Prize. This is one prize that is well known to even the lay public. This prize is given every year to deserving candidates in each of these six categories.

Five of these prizes, namely those for physics, chemistry, medicine, literature, and peace, were instituted by Alfred Nobel who was born in Sweden in 1833. He was a chemist, an engineer and an inventor and held 355 patents. Among his inventions, dynamite is the most famous. It is an explosive which was of great use in the mining industry, construction works, rail road construction etc. Nobel made a lot of money from this and was extremely wealthy. Dynamite was unfortunately also used extensively in weapons used by armies, and caused large scale devastation and a large number of deaths. This earned notoriety for Alfred Nobel.

According to a well known story, when his brother died in 1888, a French newspaper, mistakenly assuming that Alfred Nobel had died, published an obituary calling him a "Merchant of death". This deeply saddened Alfred and he decided to use his wealth to establish the Nobel Prizes which were to be given to researchers/writers/workers in the five fields mentioned above, "who

P. Khare, *Nobel Prizes in Astronomy*, Astronomers' Universe,
https://doi.org/10.1007/978-3-031-29639-0_1

have done the best work in that year for the benefit of mankind." The details were stated in his will which he signed on the 27th of November, 1895, about an year before his death. Nobel donated 94% of his wealth, i.e. 31 million Swedish Kronor (SEK) for establishing these prizes. The Nobel Foundation was established for the purpose, after his death. The Foundation manages the finances and administration of the Nobel money and prizes. It invests 80% of the money in shares and other investments. As per his instructions, the peace prize is given by a Norwegian Nobel committee.

These prizes have been awarded annually since 1901. The economics prize was instituted in 1968 in the memory of Alfred Nobel and is called Nobel Memorial Prize. The monetary award of each prize in 1901 was 150782 SEK and by 2020 it has risen to 10 million SEK which is close to 1.2 million US$. These prizes are not given posthumously. This condition was added only in 1974. According to this condition, the only case in which the award can be given posthumously is if the awardee was alive at the time of declaration of the award. Before 1974, two awards—the literature award for the year 1931 and the peace prize for the year 1961—were given posthumously. A prize for a particular year in a particular subject can be shared at most by three persons or organizations. If shared by two persons, the prize money is divided equally between the two. If it is shared by three persons, each of the three persons can receive one-third share of the prize or two together can receive half share, the other half going to the third person.

There have been some years when some of the Nobel Prizes were not given. These mostly were the World War years from 1914 till 1918 and 1939 till 1945. In addition some of the prizes were not given in some years as no suitable candidate could be found. There is a provision for doing so, as according to the statutes of the Nobel Foundation, "If none of the works under consideration is found to be of importance, the prize money shall be reserved until the following year. If, even then, the prize cannot be awarded, the amount shall be added to the Foundation's restricted funds." This happened in the case of Einstein. He was awarded the Nobel Prize in Physics for the year 1921, but it was announced in 1922 as no work was found suitable in the year 1921! Of course, Einstein had done his Nobel Prize winning work in 1905 itself.

Even though Nobel in his will instructed that the prizes for a particular year were to be given to researchers in the five fields mentioned above, who have done the best work in that year for the benefit of mankind, this time frame has not been followed. Often it takes time for a work to be accepted or verified by peers. There are many examples of researchers being recognized several years after their path breaking work was completed. An example is that of Chandrasekhar Subrahmanyan who was awarded the prize in 1983 for his

work on the limit on the masses of White Dwarf stars which he had completed in the 1930s. There are several other examples as well.

Box 1.1 No Prize for Mathematics

Surprisingly, Nobel did not include a prize for mathematics, which is considered by many to be the queen of the sciences. Perhaps Nobel did not consider it as a science useful for humanity. This led famous Norwegian mathematician Sophus Lie to propose a similar prize for mathematics just before the end of the nineteenth century. It was to be given in the name of another famous Norwegian mathematician Niels Henrik Abel (1802–1829) and was to be established in 1902 which was Abel's centenary year. The then king of Norway, Oscar II, agreed to finance such a prize. However, due to change in political scenario in the region, the Abel Prize was established by the Norwegian government only in 2001. The prize is more like a recognition of lifetime achievement and is given every year. Its monetary value is about US$885,000. The first Abel Prize was awarded in 2003. Perhaps a more prestigious award in the field of mathematics is the Fields Medal, given by the International Mathematical Union. It is considered to be the Nobel Prize in mathematics. The prize is given every 4 years to a minimum of two and a maximum of four mathematicians who are less than 40 years old. The medal is named after Canadian mathematician John Charles Fields, who was instrumental in establishing the award and also gave monetary help for its establishment. The first such medal was awarded in 1936. The monetary value of the award is about 15,000 Canadian Dollars.

The youngest recipient of the Nobel Prizes so far has been Malala Yousafzai, who won the Nobel Peace Prize for the year 2014. She was just 17 years old then. The oldest recipient of the Nobel Prize so far has been John B. Goodenough, who won the Nobel Prize in Chemistry for the year 2019 at the age of 97.

So far only four individuals have won the Nobel Prize twice. These are, M. Curie (physics 1903, chemistry 1911), J. Bardeen (physics 1956 and 1972), L. Pauling (chemistry 1954, peace 1962) and F. Sanger (chemistry 1958 and 1980). There also have been several family members to have won the prize, the most prolific of them being the Curies. Marie Curie and her husband Pierre Curie won the Nobel Prize for physics in 1903, Marie Curie again received the prize for chemistry for the year 1911, and her daughter Irene Joliot-Curie along with her husband Frederik Joliot received the Nobel Prize for chemistry in 1935. There have been six cases of a father-son duo who were awarded the prize, out of which only one pair was awarded for the same subject in the same year, which was the physics prize in 1915 to William Henry Bragg and William Lawrence Bragg. Till date, there have been 58 women Nobel Prize winners. In

the rest of this chapter, some technical terms have been used without defining them. They will be defined wherever they appear later in the book.

In the first half of the nineteenth century, astronomy was considered to be purely an observational science. It mainly dealt with positions of stars and positions and motion/orbits of planets. In the second half of that century, with the development of optical instruments like diffraction grating and the spectrograph, the spectroscopic observations of stars and planets were undertaken and it was discovered that the dark lines in the spectrum of the Sun were produced by the same chemical elements that existed on the Earth. Soon the connection between astronomy and physics started attracting attention. By the end of the century, the connection was firmly established and a new branch, namely astrophysics—which deals with the physics of astronomical objects—gained recognition. In spite of this recognition, astronomers were held in low esteem by the physicists who often ridiculed their science as being accurate only up to an order of magnitude.

With new electronic detectors, bigger telescopes and development of space technology which enabled observations in other bands of electromagnetic radiation (described in Box 2.1), the accuracy of astronomical measurements improved rapidly with time and by the late twentieth century, measurements of a number of astronomical quantities, accurate up to several decimal places, could be made. Finally, physicists started taking astronomers seriously and their work slowly started receiving the highest recognition.

Box 1.2 Physics and Astronomy

There are several instances of physics having gained from discoveries in astronomy. Newton could not have discovered gravity and its laws without the observation of planetary motions. Helium was first discovered in 1868, in the Sun, about 27 years before it was discovered on the Earth. Astronomy contributed significantly to the formulation of Einstein's general theory of relativity. Confirmation of this theory also used astronomical observations. Studies in spectroscopy and nuclear fusion were boosted by astronomical considerations and observations. Some properties of elementary or fundamental particles called neutrinos were discovered mostly because of the solar neutrino problem, about which we will read in Chap. 2. These are just a few examples of astronomy leading the way for physics.

There are also several spin-offs from astronomy which have helped and enriched our technology. From the charge coupled detectors to some astronomical software which is of use in the car industry, computer systems, petrochemical industry etc., there are plenty of examples.

In what follows, the word astronomy is used in the broad sense of the term to include observational astronomy: observations of the position, motion and spectroscopic properties of astronomical objects, astrophysics: the physics of astronomical objects, and cosmology: the study of the origin and evolution of the Universe as a whole.

Before 1974, four Nobel Prizes for physics had been awarded to physicists working in subjects which were relevant to astronomy. These were for years (i) 1936, to Victor Franz Hess for his discovery of cosmic radiation, (ii) 1947, to Edward Victor Appleton for his investigations of the physics of the upper atmosphere, especially for the discovery of the so called Appleton layer, (iii) 1967, to Hans Albrecht Bethe for his contributions to the theory of nuclear reactions, especially his discoveries concerning the energy production in stars, and (iv) 1970, to Hannes Olaf Gösta Alfvén for his fundamental work and discoveries in magneto-hydro-dynamics with fruitful applications in different parts of plasma physics. Out of these, only the 1967 award to the nuclear physicist, Bethe is included in this book, as his work is of fundamental importance for understanding the structure of stars. The first Nobel Prize for physics awarded to professional astronomers was for the year 1974 when it was given to radio astronomers Martin Ryle and Anthony Hewish. So far astronomers have been awarded prizes in 10 years. Out of these, 6 have been in this century itself.

The aim of this book is to describe the Nobel Prize winning work of astronomers to all readers having a background of high school physics. Obviously the work, being at the forefront of astronomy, involves advanced physics and mathematics. However, an attempt has been made to avoid the mathematics completely, and instead, explain the physics behind the work in a language which should be understandable even to a reader with a high school level science background. This should give an idea of the work that has been (and is being) done at the forefront of astronomy, specially to high school and college students and to all other curious minds having an interest in learning more about our Universe. It is hoped that this attempt is successful and the readers feel motivated to read more advanced material on the subject, and also that at least some of the students get sufficiently interested and study the subject during their college years, and who knows, be future Nobel laureates themselves.

In the rest of the book, the works of Nobel Prize winning astronomers are described. Table 1.1 gives the list of Nobel Prizes awarded in the field of astronomy. Along with the year of award and the names of the astronomers, it also gives the citation for the Nobel Prizes. The Nobel Prize winners in Table 1.1 have been collectively referred to as astronomers, keeping in mind that Bethe

was not one, and was a nuclear physicist. In Table 1.1 the Nobel Prizes awarded to astronomers are ordered according to the year of the award. The awards were given for 11 years (ten years to astronomers and one year to Bethe) and were won by 26 astronomers. They covered 16 different subjects. The awards can roughly be divided in seven categories depending on the subject they are awarded for. These, along with the years of award, are

1. Stellar structure (1967, 2002)
2. Stellar evolution (1983)
3. Radio and X-ray astronomy (1974, 2002)
4. Extrasolar planets (2019)
5. Black holes (2020)
6. Gravitational waves (1993, 2017)
7. Cosmology (1978, 2006, 2011, 2019)

These categories have also been included in Table 1.1.

The works in the seven categories are described in the seven chapters that follow. The chapters have been ordered keeping in mind the ease of understanding for the reader. The first section in each chapter introduces the contents of that chapter. After that, each section describes the Nobel Prize given for one of the subjects in that category. Each section is divided into three parts. The first part gives the citation and short biographical information about the awardee(s). The second part gives the background necessary to understand the Nobel Prize winning work and the third part describes the actual prize winning work. The title given to this section describes the main subject of the work. The order of presentation of the prizes in each chapter is chosen so as to make understanding the works easier.

Table 1.1 List of Nobel Prizes in astronomy

Year	Name	Citation and (category)
1967	Hans A. Bethe	Contributions to the theory of nuclear reactions, especially his discoveries concerning the energy production in stars (1)
1974	Martin Ryle	Pioneering research in radio astrophysics: observations and inventions, in particular of the aperture synthesis technique (3)
	Antony Hewish	Pioneering research in radio astrophysics: for his decisive role in the discovery of pulsars (3)
1978	Arno Penzias, Robert W. Wilson	Discovery of cosmic microwave background radiation, providing support for the big bang theory (7)
1983	Subrahmanyan Chandrasekhar	Theoretical studies of the physical processes of importance to the structure and evolution of the stars (2)
	William A. Fowler	Theoretical and experimental studies of the nuclear reactions of importance in the formation of the chemical elements in the Universe (2)
1993	Russell Alan Hulse, Joseph H. Taylor Jr.	Discovery of a new type of pulsar, a discovery that has opened up new possibilities for the study of gravitation (6)
2002	Raymond Davis Jr., Masatoshi Koshiba	Pioneering contributions to astrophysics, in particular for the detection of cosmic neutrinos (1)
	Riccardo Giacconi	For pioneering contributions to astrophysics, which have led to the discovery of cosmic X-ray sources (3)
2006	John C. Mather, George F. Smoot	Discovery of the black body form and anisotropy of the cosmic microwave background radiation (7)
2011	Saul Perlmutter, Brian P. Schmidt, Adam G. Riess	Discovery of the accelerating expansion of the Universe through observations of distant supernovae (7)
2017	Barry C. Barish, Kip S. Thorne, Rainer Weiss	Decisive contributions to the Laser Interferometer Gravitational-Wave Observatory detector and the observation of gravitational waves (6)
2019	James Peebles	Theoretical discoveries in physical cosmology (7)
	Michel Mayor, Didier Queloz	Discovery of an exoplanet orbiting a solar type star (4)
2020	Reinhard Genzel, Andrea M. Ghez	Discovery of a supermassive compact object at the centre of the Milky Way Galaxy (5)
	Roger Penrose	Discovery that black hole formation is a robust prediction of the general theory of relativity (5)

2

Stellar Structure

2.1 Introduction

The Nobel Prizes in the category of stellar structure were awarded in years 1967 and 2002 to three astronomers: Hans Albrecht Bethe, Raymond Davis Jr. and Masatoshi Koshiba. The award winning work for the year 1967 by Bethe is purely theoretical and is mainly in the field of nuclear physics as applied to astrophysics. It involves the energy sources in the stars and was done in the late 1930s. The work for the award of 2002 is in the field of experimental particle physics and astronomy. It involves detection of elementary particles called electron neutrinos, which are produced inside the Sun and in exploding stars, and are able to reach the Earth. The experimental work for this award was initiated by Davis in 1965. Davis announced his results in 1968. Koshiba's work started much later, after 1983. He detected neutrinos coming from a supernova explosion in the nearby galaxy Large Magellanic Cloud in 1987, and later also detected those coming from the Sun. We will be learning about supernova explosions in the next chapter. Koshiba's results about solar neutrinos were published in 1989 and they reaffirmed Davis's results.

P. Khare, *Nobel Prizes in Astronomy*, Astronomers' Universe,
https://doi.org/10.1007/978-3-031-29639-0_2

2.2 Nobel Prize 1967: H. A. Bethe

The Nobel Prize for the year 1967 was given to Hans Albrecht Bethe "for his contributions to the theory of nuclear reactions, especially his discoveries concerning the energy production in stars." He was the sole recipient of the award.

Credit: By Los Alamos National Laboratory—Los Alamos National Laboratory, Attribution, https://commons.wikimedia.org/w/index.php?curid=71763

Hans A. Bethe was born in Germany in 1906. He did his Ph.D. under the famous German physicist Arnold Sommerfeld at a very young age of 22 years. He then taught at different places in Germany. In 1933 he lost his job owing to the Nazi regime and migrated to England and later, in 1935 to the USA, where he joined as a faculty in Cornell University. He continued there till the end except for his work at Los Alamos and during his sabbaticals. He was one of the greatest nuclear physicists of the twentieth century. He was also a great human being and was highly admired for his scientific achievements, his integrity, fairness, and for his deeply felt concern for the progress of science and humanity.

In war time, Bethe joined the Manhattan project and made important contributions towards the making of the first atom bomb. Later he also contributed, though reluctantly, towards the making of the hydrogen bomb. He was one of the first members of the President's Scientific Advisory Committee formed in 1957 and he worked towards a test ban treaty and towards constraining further development of more powerful weapons. He persuaded two American presidents, John F. Kennedy and Richard Nixon, to sign Partial Nuclear Treaty and Anti-Ballistic Missile Treaty in 1963 and 1972 respectively.

Bethe was a mainly a nuclear physicist and worked mostly on the structure of nuclei. He made several fundamental contributions in the field. In addition

he has done major works in atomic physics and even in solid state physics. He continued making important contributions in his nineties also. After contributing to scientific research at the highest level for seven decades, he passed away in 2005.

2.2.1 Background

About 4000 stars are visible to the naked eye in the night sky. All astronomical objects in the sky appear to be moving from east to west, which is due to the rotation of the Earth around itself. Apart from this apparent motion, the positions of the planets in the sky keep changing over a period of days, while stars seem to remain at fixed points in the sky. However, it was discovered even in ancient times that stars do move a little, when observed over many years. This led to the idea that stars were much farther away than the planets. We now know that the stars we see with our eyes in the night sky are a very tiny fraction of about a hundred billion stars that constitute the Milky Way galaxy. The Milky Way is a spiral galaxy and has a disc like structure with stars mostly concentrated along spiral arms in the disc. The disc has a diameter of about a hundred thousand light years, a light year being the distance traveled by light in one year; it is equal to 9.641×10^{12} km. Most of the gas and dust in the Galaxy (the word Galaxy with an upper case G stands for The Milky Way) is also present in the disc.

Over the past hundred years or so, using large optical and radio telescopes, as well as space based telescopes working in other bands of the electromagnetic radiation (see Box 2.1), along with sophisticated instruments, e.g., electronic detectors called the Charge Coupled Devices (CCDs), high resolution spectrographs etc., astronomers have been able to measure the distances, masses, sizes, temperatures, chemical compositions or abundances (which are the relative amounts of different chemical elements in stellar material) of a very large number of stars. Applying the laws of physics to these data, astronomers have developed a very good, though not yet complete, understanding of the nature of stars, how they form, what is their structure, how they generate energy, how long they will live, and what happens to them at the end of their life cycle. We will take a brief look at the structure of the stars below.

Box 2.1 Electromagnetic Radiation

The light that we can see, i.e., to which our eyes are sensitive, is a wave called electromagnetic wave. In any wave, the magnitude of some particular entity goes through periodic increase and decrease; for example in the case of a water wave, this entity is the vertical distance of a water particle above or below the average surface level of water. A typical wave is depicted in Fig. A where the value of the varying entity is plotted along the vertical axis as a function of distance along the direction of travel of the wave (which is plotted along the horizontal axis). The maximum value of the varying entity, e.g., OA, is called the amplitude of the wave. The distance between two consecutive points in the same state of vibration, e.g., A and C, B and D, or X and Y, along the wave is called its wavelength. It is the distance over which the wave pattern repeats itself. The frequency of a wave is inversely proportional to the wavelength, the product of the two being equal to the velocity or speed of the wave. A plot of the value of the varying entity at any given space point as a function of time will be identical to Fig. A. For an electromagnetic wave, the varying entities are the strengths of the electric and the magnetic field. A wave is characterized by its wavelength or frequency.

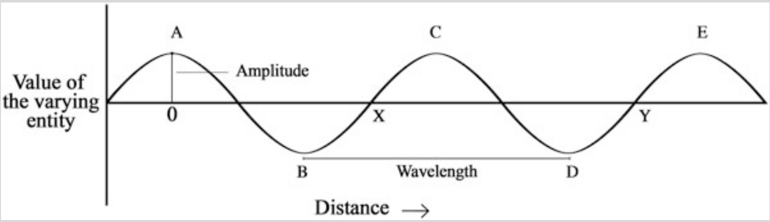


Fig. A

Electromagnetic radiation, or electromagnetic waves, can have an infinite range of wavelengths starting from very small, smaller than a nanometre (nm, which is equal to 10^{-9} m) to larger than a kilometre (km). The electromagnetic waves to which our eyes are sensitive are called visible light/radiation and have wavelengths between 400 to 700 nm. The electromagnetic radiation has been divided into 7 bands, which, in increasing order of wavelength, are called gamma rays, X-rays, ultraviolet, visible, infrared, microwave and radio waves. These are shown in Fig. B. The coloured portion at the centre shows the visible radiation. The wavelength is plotted in metres (m).

Electromagnetic radiation has dual nature; it can behave like a wave, as seen above, but can also behave like a particle. These particles are called photons. The energy of the photons decreases with increasing wavelength, with gamma ray photons having the maximum energy and radio photons having the minimum energy. Some physical phenomena involving electromagnetic radiation, e.g., interference (described in Chap. 4), can only be explained using the wave nature

(continued)

Box 2.1 (continued)

of light, while some other phenomena, e.g., the radiation emitted by atoms and the photoelectric effect require the particle nature of light for their explanation.

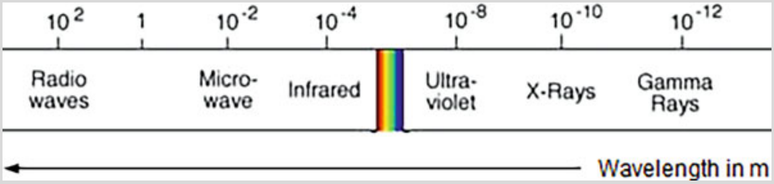


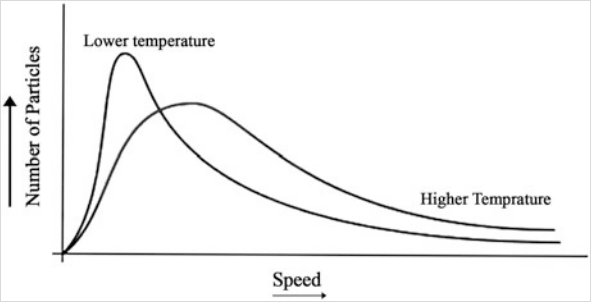
Fig. B

Astronomical sources emit electromagnetic radiation of some or all of these wavelengths. The Earth's atmosphere, however, completely absorbs all waves except the visible radiation and the radio waves. These two types of waves do undergo absorption by the atmospheric gases but the absorption is not severe. Thus, we can observe astronomical sources in visible light and in radio waves from the surface of the Earth.

A star is a very large sphere of gas which emits a huge amount of energy. The stars in the night sky appear like point objects to us only because they are at very large distances from us. The star closest to us is the Sun; we depend on it for our existence. The Sun has a mass of 2×10^{30} kg, a radius of 7×10^5 km and is at a distance of about 150 million km from the Earth. It is about 267,000 times closer than the next closest star, Proxima Centauri which is at a distance of about 4×10^{13} km from us. The rate of emission of energy of a star is called its luminosity. The luminosity of the Sun is about 4×10^{33} ergs s^{-1}, erg being a unit of energy. The yellowish colour of the Sun tells us (see Box 2.3) that the temperature at the surface of the Sun is about 5800 K, K indicating the Kelvin scale. The Kelvin scale is similar to the Celsius scale except that its zero occurs at −273.15 °C. So in this scale, ice melts at 273.15 K while water boils at 373.15 K, and 0 K is the minimum temperature possible anywhere in the Universe. The material from which the Sun had formed contained about 72% by mass of hydrogen, about 26% of helium and the rest 2% consisted of heavier elements like carbon, nitrogen, oxygen etc., all in gaseous form.

Box 2.2 Classical Ideal Gas

A gas whose particles occupy negligible space and do not interact with one another is called an ideal gas. Though the particles of a real gas, for example, the gas present in our atmosphere or that present inside the Sun and most other stars, do interact with one another, the force of interaction is usually extremely small and can be ignored. These gases can, therefore, be considered to be ideal gases. There are huge gas clouds present in the space in between stars, called the interstellar clouds, which have extremely low density and the gas can be again considered to be ideal. Exceptions to ideal behaviour of gas occur in gases having extremely high density at the centres of certain stars which will be discussed in the next chapter. For now, we will only consider an ideal gas.



The particles of a gas are in constant random motion. The particles of ideal gases mentioned above, obey classical (Newton's) laws of motion, and these gases are called classical ideal gases. The speeds of particles of a classical ideal gas have a particular distribution known as the Maxwell–Boltzmann distribution. A plot of the number of particles of the gas as a function of speed for a classical ideal gas is shown in the figure for two different temperatures. It is clear from the figure that the most probable speed, i.e., the speed which maximum number of particles have, or the speed at the peak of the distribution, as well as the average speed of the particles (not shown in the figure) increases with increase in temperature of the gas.

We know that gas exerts pressure. This can be seen from the fact that when gas/air is inserted in a balloon, the balloon blows up. The sides of the balloon are pushed outwards due to the pressure of the inserted gas. The pressure exerted by a gas on any given surface is the force exerted by gas particles on a unit area of the surface. This is determined by the number of collisions that the particles have per unit area of the surface per unit time and on the energy of the particles. Both these quantities increase with increase in the temperature as well as in the particle density (the number of particles per unit volume) of the gas. This is the reason why gas pressure increases with increase in the temperature and density of the gas. This fact can also be understood from the example of the balloon: when a balloon is kept inside a refrigerator, it shrinks because of the lower pressure of

(continued)

Box 2.2 (continued)
the now cooler gas inside it. Also, the higher the density of the gas, the higher is the pressure: when we blow more air into the balloon, we increase the density of the gas and the balloon expands due to the higher pressure of the gas inside it.

The Sun is a typical or average star in the Galaxy, in the sense that there are many stars which are (i) more as well as less massive than the Sun, (ii) hotter as well as cooler than the Sun, and (iii) bigger as well as smaller in size as compared to the Sun. We can therefore use the properties of the Sun as a unit: the masses of stars are stated as multiples or fractions of the solar mass ($M_\odot$) and the radii of stars are stated as multiples or fractions of the solar radius ($R_\odot$). In these solar units, most stars have masses in the range of about 0.1 $M_\odot$ to a few hundred $M_\odot$. The radii of most stars are in the range of a fraction of $R_\odot$ to about 1000 $R_\odot$. The present age of the Sun is about 4.5 billion years, and from the theory of stellar evolution that we will learn in the next chapter, we expect that it will have a total lifetime of about 9 billion years. For most stars the lifetime varies from a few million to around ten billion years. The higher the mass of a star, the shorter is its total lifetime. The Sun being the star closest to us, we have been able to study it in great detail. We have been able to measure the magnetic field on the surface of the Sun and study the oscillations of its surface layers, and through that, the vibrations of its interior as well. We also have been able to look directly at the centre of the Sun, as we will see later in this chapter. Below, the structure of the stars is discussed, often referring to the Sun in particular, keeping in mind that the discussion is applicable to other stars as well.

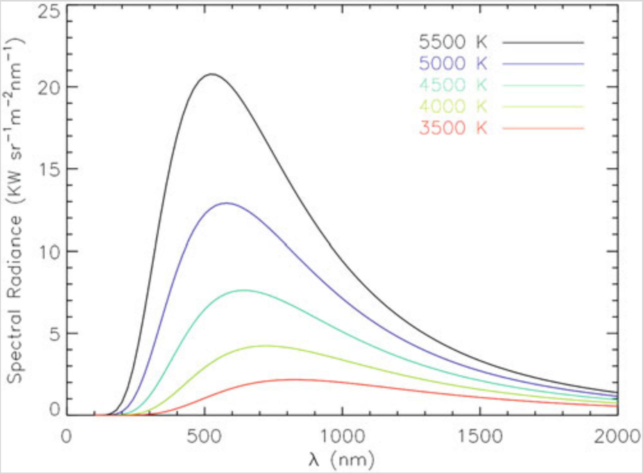
Box 2.3 Thermal and Black Body Radiation

Every object having a temperature above absolute zero (0 K) possesses heat energy, i.e., thermal energy and emits radiation because of this energy. The radiation emitted by an object on account of its thermal energy is called thermal radiation. This radiation is in the form of electromagnetic waves of different wavelengths. The amount of thermal radiation emitted by a body at different wavelengths depends on the properties of the body. However, there exists a class of objects called black bodies which have a characteristic spectrum, i.e., the distribution of the amounts of the radiation emitted as a function of frequency or wavelength. The shape of the quantitative spectra of these bodies, i.e., a plot

(continued)

Box 2.3 (continued)

of the amount of the radiation emitted as a function of frequency or wavelength, depends only on the temperature of the body. The total amount of energy emitted by a black body depends also on its surface area.



The amount of energy emitted by a black body at different wavelengths, for five different temperatures of the black body, is shown in the figure. Each curve is known as a black body spectrum. The total amount of energy emitted by the body is equal to the area under the curve. It can be seen that the higher the temperature of the black body, the higher is the total amount of radiation emitted, and the lower is the wavelength at which maximum energy is emitted. We actually see this in our daily life. Remember that an iron rod when heated changes its colour. As its temperature rises, its colour changes in the following order: black, red, orange, yellow, white, and bluish white. This is because, the rod actually behaves like a black body. As the temperature of the rod increases, it emits more and more energy in radiation of smaller and smaller wavelength. This is responsible for the change in its colour perceived by us.

An ideal black body is one which absorbs all radiation falling on it and does not allow any radiation to go out. The radiation inside it is characteristic of the temperature of the black body. In practice, a body which does emit radiation, but for which the amount of radiation emitted is much smaller than the radiation trapped inside it, can be approximated to be a black body. Stars do fulfill this criterion and the radiation coming out of them is characteristic of a black body having the temperature of their surface. The luminosity of a star is determined by its surface area, i.e., its radius, and its surface temperature.

Equilibrium of Stars

Geological studies of fossils, and radioactive materials in rocks on the Earth, on the Moon, and the material in meteorites, along with the results of theoretical astronomical studies, indicate that the properties of the Sun, i.e., its temperature and luminosity, could not have changed significantly over the past 4.5 billion years. The luminosity of a star is determined by its temperature and its size (see Box 2.3). Thus, constant luminosity and temperature also means that the size of the Sun could not have changed over this time period. As per our current theoretical understanding of stars, these properties of the Sun will remain almost unchanged in the future for a similar period of time. Why is it that the Sun, which is a sphere of hot gases, will remain in the same state for nearly 9 billion years? The reason is that it is in a near-perfect mechanical and thermal equilibrium state.

There are two competing forces that have opposite effects on the structure of every star. One is the familiar attractive force of gravity which tries to pull all the matter of the star closer together towards one point, its centre. Due to the large mass of the star, this force is so strong that if it were unopposed, the star would collapse to a point in a couple of hours. Obviously, for a star to maintain its size for billions of years, the force of gravity needs to be opposed and exactly balanced by another force. We have seen (see Box 2.2) that gas exerts outward pressure which increases with its temperature and density. Due to this pressure, the hot gas comprising a star tries to disperse into the near-vacuum (nearly zero pressure) conditions of the surrounding space, and thus the pressure provides an outward force, opposing gravity. For stars like the Sun, both these opposing forces exactly balance each other at every point inside it. This mechanical equilibrium state of the stars is known as the state of hydrostatic equilibrium.

The balance of forces i.e., hydrostatic equilibrium of the star can also be looked at in another way. The pressure at any depth inside the star needs to be high enough to support the weight of all the matter above it. As we go deeper into the star, the height of the outward layer and therefore, the weight to be supported is greater, so the pressure has to be higher. This is similar to the case of a person diving in a water body on the surface of the Earth. The lower she goes below the surface of water, the higher is the weight of the water layer above her and the higher is the pressure she feels. Thus, the pressure in a star increases as the depth from its surface increases. This is shown in Fig. 2.1. Increasing pressure with depth in turn implies that the density and the temperature of the gas must increase as we go closer to the centre. Using the equations of physics which describe these forces and their balance, it can

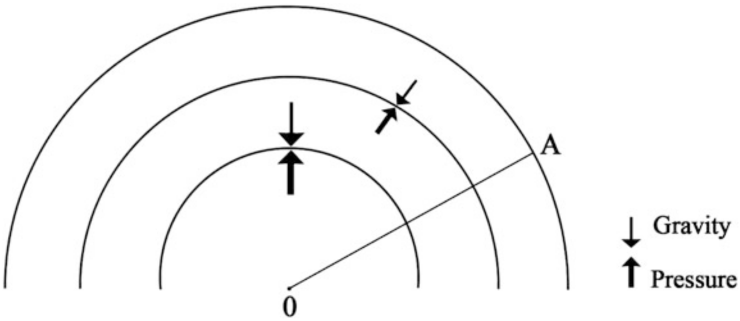


Fig. 2.1 Figure showing the forces of gravitation, shown by thin arrows, acting inwards and those of gas pressure shown by thick arrows, acting outwards at different distances from the centre O of a star. The lengths of the arrows indicate the magnitude of the forces. Only a part of the star has been shown in the figure. OA is the radius of the star

be shown that the temperature at the centre of the Sun is about 20 million K, which is much higher than the temperature at the surface which is about 5800 K. The larger the mass of a star, the higher is its central pressure as it has to support a larger weight. This needs higher temperatures.

Energy Generation in Stars

We know that heat always flows from a hotter body to a cooler body. The Sun, being much hotter than the surrounding cold space, is constantly losing heat, i.e., thermal energy, to the surrounding in the form of radiation. We know this very well as all the energy necessary for our survival comes from the Sun in the form of radiation. Loss of heat should cause the material of the Sun to cool, resulting in a decrease in its gas pressure. On the other hand, the force of gravity in the Sun has not changed with time, because, as we saw above, its properties have not changed over past 4.5 billion years. This means that the equilibrium between gas pressure and gravity has been maintained all through its lifetime. For this to happen, the gas pressure and hence, the temperature should have remained constant. It is therefore, essential that the heat lost by the Sun in the form of radiation be constantly replenished. Thus, we come to the conclusion that a source of energy must be present inside the Sun which constantly supplies energy and maintains its temperature. We know that the luminosity of the Sun, i.e., the amount of energy that it is giving out per second is about 4×10^{33} erg s^{-1}. In its lifetime so far, of 4.5 billion years, i.e., roughly 1.4×10^{17} seconds it has radiated a total of 5.6×10^{50} ergs of energy. The mass

of the Sun is about 2×10^{30} kg and so it has emitted 2.8×10^{20} ergs per unit kg of its mass till now.

What could be the source which can supply this much energy per kg? Early on, various possibilities were considered. One of them was that the Sun has no source of energy but was created very hot and has been constantly losing its heat energy, i.e., its thermal energy ever since its creation and has reached the present temperatures at its present age. Another possibility that was considered was that the Sun has been slowly contracting and its gravitational potential energy (see Box 2.4), which is negative and is inversely proportional to its radius, is decreasing, i.e., becoming more and more negative with time (as the radius is decreasing due to contraction). We know that energy is always conserved. Thus, the lost gravitational potential energy could have been the energy radiated by the Sun. It can be shown, using simple equations of physics, that both these types of energies, i.e., the thermal energy and the gravitational potential energy of a star, are related to each other and that half of the decrease in the gravitational potential energy would be stored as thermal energy in the Sun. The other half would be emitted as radiation. Thus, both the above possibilities are essentially the same. We can calculate the present gravitational potential energy of the Sun from its mass and radius. This is also the total decrease in its gravitational potential energy since its birth, as this energy was zero when the Sun was in the form of a tenuous gas cloud (stars are believed to form from contraction of very tenuous, i.e., rarefied interstellar clouds). Thus, the total amount of energy emitted by the Sun in its lifetime using its gravitational potential energy can be calculated from its mass and radius, and it comes out to be 3.8×10^{18} erg per kg which is only 1.3% of the energy per kg emitted by the Sun so far. So these sources of energy generation in the Sun can be easily ruled out. It is also ruled out from the fact that for this to be the source of energy, the radius of the Sun would have been constantly decreasing, which we know to be not correct.

Box 2.4 Gravitational Potential Energy

Potential energy is the energy that an object possesses because of its position. We can understand it easily from the example of a ball. Let us consider a ball thrown vertically upwards from point A on the ground. It goes up to a maximum height, up to point B say. At A its speed is maximum and at B it is zero. After that it starts falling back to ground. Let C be a point midway between A and B along the path of the ball. On reaching the Earth it probably bounces a few times and comes to rest. Now let us consider the energy of the ball at the three points A, B and C.

(continued)

Box 2.4 (continued)

At A its kinetic energy (energy of motion) is the highest as it has highest speed there, and at B it is zero. At C it is in between the two values.

We know that energy is always conserved. So what happened to the kinetic energy of the ball as it went from A to B? Well, the ball has another type of energy, gravitational potential energy, due to its position with respect to the Earth. This energy is minimum at A and maximum at B. At C it is in between the two values. From where did the ball acquire this gravitational potential energy? It came from the kinetic energy of the ball, which in turn came from the energy that we spent for throwing the ball upwards. Thus, during the motion of the ball upwards, its kinetic energy gets converted into gravitational potential energy. On the way back, the opposite conversion takes place and the ball gains kinetic energy at the expense of its potential energy.

What happens when the ball falls on the ground? Its energy gets lost in the form of heat due to friction with the ground, and also in the form of any sound that may be generated, and ultimately it loses all its energy and comes to rest. The potential energy of the ball arose because of the force of gravity exerted on it by the Earth. There are other forms of potential energy which come into play due to other types of forces acting on a body. For example a charged particle in an electric field experiences electric force and will have electrostatic potential energy.

The gravitational potential energy of a star is negative, and its magnitude is equal to the work done by gravity to form the star from material which is distributed over a huge interstellar cloud and can be assumed to be distributed over infinite volume. As the gas cloud contracts, the gas particles come closer and closer, the work being performed by gravity. As gravity is actually bringing the particles closer under its influence, the work done by it is positive and the gravitational potential energy of a star is negative. Its magnitude is directly proportional to the mass of the star and inversely proportional to the radius of the star. Thus, as a star contracts, i.e., its radius becomes smaller, the magnitude of the potential energy increases, but being a negative quantity, the energy actually decreases.

The third possibility that was considered for the source of energy in the Sun, was chemical energy which may be getting converted into heat and radiation through chemical reactions. Chemical reactions convert one or more chemicals (atoms and molecules) into other chemicals, for example the burning of coal in which carbon and oxygen combine to form carbon dioxide, or the combination of hydrogen and oxygen to form water. However, the chemical energy that can be released per kg of any material is only about 10^{14} erg and falls far short of the energy released by the Sun per kg so far. Thus, chemical reactions can also be ruled out as a source of energy inside the Sun.

After ruling out the above three sources of energy, scientists reached the conclusion that the only possible source of energy in the Sun and other stars

must be nuclear energy. This energy is released in some types of nuclear reactions, which are processes which convert one type of nucleus, e.g., carbon, into another type of nucleus, e.g., oxygen. According to Einstein's famous theory of special relativity (see Box 2.5), energy and mass are equivalent and mass can be converted into an equivalent amount of energy. Nuclear reactions, as we will see below, do just that. The amount of energy released by converting one kg of solar mass into energy is 1500 times that emitted by Sun per kg since its birth. So nuclear reactions are a viable source of energy in the Sun. Let us learn more about these reactions.

Box 2.5 Special Theory of Relativity

The special theory of relativity is a theory of the relationship between space, time and motion. It was developed by Albert Einstein and was presented in 1905. Before Einstein gave this theory, Newton's laws of motion given in 1686 were used for studying the motion of objects. These laws, however, were found to be inconsistent with the behaviour of electromagnetic waves. The paradox was solved by Einstein through his special theory. Important outcomes of this theory are (1) the velocity of light is constant and does not depend on the velocity of the source or the observer, (2) no object or information can travel faster than the velocity of light in vacuum, (3) mass and energy are interchangeable, and the energy equivalent to a given mass is equal to the product of that mass and the square of the velocity of light in vacuum, (4) a particle with zero mass is bound to move with the velocity of light and (5) three dimensional space and time form a four dimensional entity called space-time; space and time are entangled with each other.

Some interesting consequences of the special theory are (i) a clock moving with respect to an observer appears to go slower than a clock which is at rest with respect to her; (ii) a rod moving with respect to an observer appears to shrink in length when compared to an identical rod at rest with respect to her; (iii) two events which appear to occur at the same time, i.e., simultaneously to an observer A, may not appear to be simultaneous to another observer B moving with respect to A and (iv) the mass of a particle in motion is greater than its mass when it is at rest.

Fundamental Forces of Nature

How one body influences the motion of another body depends upon the forces they exert on each other. For example, the Earth goes round the Sun because of the gravitational force of attraction between them. Electrically charged bodies exert an electric force on each other. How many different types of forces exist in nature? After centuries of study, physicists have established that there are only

four basic or fundamental forces in nature. These forces, in order of increasing strength, are (1) Gravitational force, (2) Weak force, (3) Electromagnetic force, and (4) Strong or Nuclear force.

Gravitational force was discovered and mathematically formulated by Newton in 1687, while electric force, which was known for a long time, was formulated by Coulomb in 1785. Magnetic force has been known at least since the sixth century, but was quantitatively understood only in the early nineteenth century, thanks to the efforts of a number of physicists. Maxwell in 1865 showed that the electric and magnetic forces are not independent. Together they are called the electromagnetic force. Gravitational force acts between any two bodies and is always attractive. On the other hand, electric force can be either attractive or repulsive. Two types of charges occur in nature. These are called positive and negative charges. Like charges, i.e., both positive or both negative, repel each other, while unlike charges, i.e., one positive and one negative, attract each other. Both gravitational and electromagnetic force have infinite range, i.e., they act even when the two bodies are at a very large distance from each other. For example, we know that the gravitational force acts between the Sun and the planets which are at distances larger than millions of km from the Sun.

The weak and the strong forces were discovered only in the twentieth century. Unlike gravitational and electromagnetic forces which act at all distances, the weak and strong forces act only over very short distances, around 10^{-13} cm (one tenth of one millionth, millionth of a centimetre), i.e., they are short range forces. Over these short distances though, these forces are much stronger than gravitational force and strong force is even stronger than electromagnetic force. The weak and the strong forces are extremely important in phenomena related to particles like protons, neutrons, etc., and to atomic nuclei, for example in nuclear reactions. They play a central role in nuclear energy generation in nuclear reactors, in nuclear weapons, and also inside stars. The strengths of the four forces, from the strongest to the weakest, are in the ratio of $1 : 10^{-2} : 10^{-13} : 10^{-38}$.

Nuclear Energy and Nuclear Reactions

An atom consists of a positively charged nucleus around which negatively charged electrons revolve. The atom as a whole is electrically neutral. The nucleus consists of protons which are positively charged particles and neutrons which are electrically neutral particles, i.e., particles having zero charge. The particles inside a nucleus, i.e., protons and neutrons are collectively called

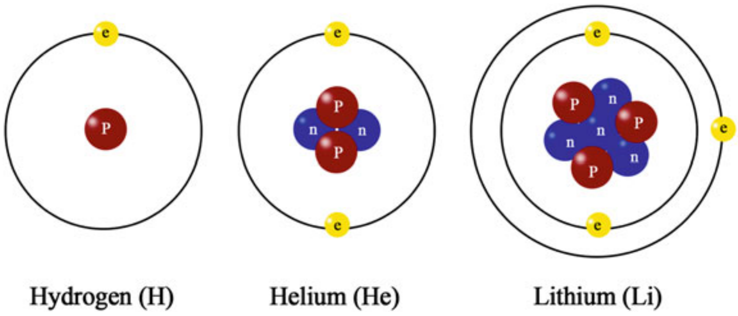


Fig. 2.2 Figure showing atoms of three of the lightest elements. Each atom has a positively charged nucleus at the centre, containing protons and neutrons, with negatively charged electrons revolving around it. Note that the figure is only illustrative and is not drawn to scale

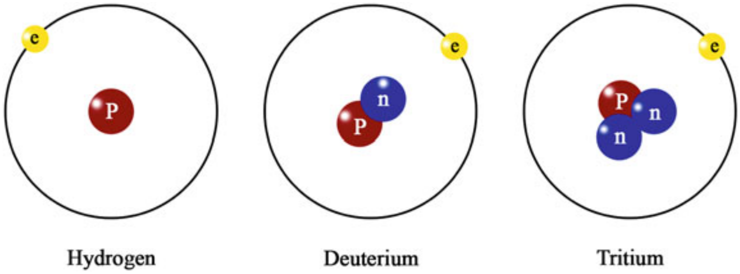


Fig. 2.3 Figure showing isotopes, deuterium and tritium of hydrogen. Note that the number of protons and electrons are same in all the isotopes

nucleons and the total number of nucleons in the nucleus of an element is called its mass number. The number of protons in the nucleus of an element is called the atomic number of that element. The atomic number is different for different elements. For example, an atom of the lightest element hydrogen has only one proton in its nucleus, while a nucleus of the next heavier element, helium, has two protons. Thus, the atomic number of hydrogen is 1 while that of helium is 2. A depiction of atoms of the three lightest elements hydrogen, helium, and lithium is shown in Fig. 2.2. Uranium, a very heavy element has an atomic number of 92, i.e., it has 92 protons. The number of neutrons in the atoms of a single element can vary by a few, and nuclei with the same number of protons but different number of neutrons are called isotopes of one another. For example a nucleus having one proton and one neutron is called deuterium, and is an isotope of hydrogen which only has one proton in its nucleus. Isotopes of hydrogen are shown in Fig. 2.3.

In a nuclear reaction, the nucleus of one element gets converted either to its isotope or to the nucleus of another element. Huge amounts of energy are released in some nuclear reactions. The reason is that in these reactions, the total mass of the products of the reaction is smaller than the total mass of the reactants. This happens due to a property of nuclei called the nuclear binding energy (see Box 2.6). So some mass is lost in the process of the reaction. This lost mass (remember the equivalence and interchangeability of mass and energy according to Einstein's special theory of relativity) is released in the form of high energy radiation or gamma rays. It may be noted that the total number of nucleons remains unchanged in a nuclear reaction.

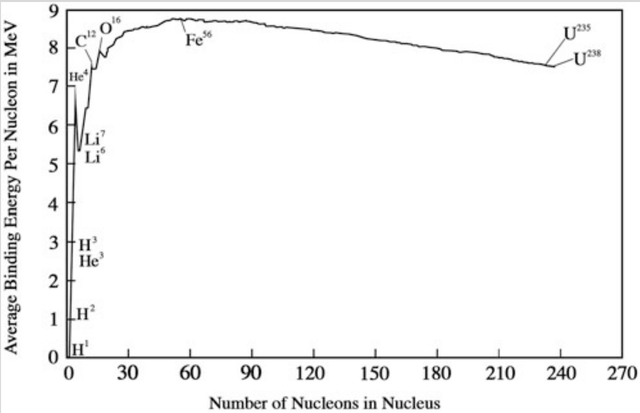
Box 2.6 Nuclear Binding Energy

The binding energy of a particle in a bound system like a planet revolving around a star or an electron in an atom, is the smallest amount of energy required to remove the particle from the system, i.e., to set it free. The binding energy of a bound system as a whole is the minimum amount of energy required to disassemble the bound system into its constituent parts. Binding energy is of different types. For example, the gravitational binding energy of a planet is the smallest amount of energy required to set the planet free from its parent star, i.e., to take it to a very large distance from the star. This is just equal to the magnitude of the gravitational potential energy of the planet. The gravitational binding energy of the entire solar system is the minimum energy required to free all planets from the Sun. The binding energy of an electron in an atom is the minimum energy required to set the electron free. The binding energy of the electron in a hydrogen atom is 13.6 eV (eV, i.e., electron volt is a unit of energy).

The binding energy of the nucleus of an element as a whole is the minimum amount of energy required to disassemble the nucleus into its constituent particles, i.e., free protons and neutrons. The typical binding energy of a nucleus is about a million times larger than the binding energy of an electron in a hydrogen atom. The higher the binding energy of a nucleus, more strongly bound the nucleus is. Considering the process which is opposite to disassembling the nucleus, binding energy is the energy which is released when a nucleus is assembled from its constituent particles. As energy is released in this process, the mass of a nucleus is smaller than the total mass of its constituent particles. For example, the mass of a helium nucleus is smaller than the combined mass of two protons and two neutrons. The binding energy of a nucleus is the energy equivalent to the difference between the total mass of the constituent particles and the mass of the nucleus as a whole. By Einstein's special theory of relativity, this equals the product of the mass difference and the square of the velocity of light in vacuum.

(continued)

Box 2.6 (continued)



For comparing how strongly nucleons are bound in nuclei of different elements, it is useful to compare the binding energy per nucleon in a nucleus. It is the ratio of the binding energy of the nucleus to the number of nucleons in the nucleus, i.e., its mass number. The higher the value of this quantity, more tightly bound are the nucleons in the nucleus. A curve showing the binding energy per nucleon versus the mass number of different nuclei is shown in the figure. It can be seen that the binding energy per nucleon increases with mass number of elements, and reaches a peak near iron-nickel-cobalt before slowly decreasing till the heaviest naturally occurring element, uranium.

There are two types of nuclear reactions that can release energy. These are fission and fusion reactions. In fission, a heavy nucleus like uranium breaks up into two or more smaller nuclei. The masses of the two product nuclei in a fission reaction are very similar. Therefore, the product nuclei will be heavier than iron nuclei. It can be seen from the binding energy curve in Box 2.6 that the binding energy per nucleon of uranium nucleus is smaller than the binding energy per nucleon of the product nuclei together, i.e., the mass of uranium nucleus is larger than the combined mass of the product nuclei. This difference in masses is emitted as radiation during the fission process. On the other hand, in a fusion reaction, a number, typically two, of nuclei of light elements combine to form a nucleus of a heavier element. As long as the product element is lighter than iron, the binding energy per nucleon of the product nucleus would be higher that the combined binding energy per nucleon of the reactants. Thus, mass of the products is again smaller than the total mass of the reactants and the difference is released as energy. Thus,

elements with masses close to that of uranium can produce energy through fission reactions, while light elements can produce energy via fusion reactions.

Which are the nuclear reactions which can possibly provide the energy for the Sun to shine? As stated above, about 72% of the mass of the Sun is made up of hydrogen. Helium is the next most abundant element making up a mass of about 26%, and the heavier elements make up only 2% of the Sun's mass. Due to the extremely small amount of very heavy elements, fission reactions seem very unlikely and fusion reactions are the most likely source of energy in stars. Among these, the most likely nuclear reaction that seems possible is the conversion of four hydrogen nuclei, i.e., protons into a helium nucleus. This process of conversion is referred to as hydrogen burning. Remember that the temperatures at the centre of the Sun and the centres of other stars are extremely high. At these temperatures, atoms of most light elements can not keep their electrons bound to their nuclei and the atoms are completely ionized, i.e., the nucleus is left by itself and the electrons roam around freely in the gas. The hydrogen and helium are, thus, present in the form of their nuclei which are protons for hydrogen and the so called alpha particles having two protons and two neutrons for helium.

For generation of energy, four protons have to come together and get bound to form a helium nucleus, converting two protons to two neutrons in the process of the nuclear reaction. The strong force is responsible for these nuclear reactions. This force acts only over very small distances of about 10^{-13} cm. Thus, for this fusion reaction to take place, four protons have to come closer than this distance. It is nearly impossible to have four protons come so close together at a given instant of time as there is a strong electrostatic force of repulsion between the protons due to their positive charges. Even if two protons have to combine, they have to have extremely high velocities for them to be able to overcome the electrostatic force of repulsion and come sufficiently close for the strong force to become operative so that the nuclear reaction can take place. We have seen that, the higher the temperature of a gas, the higher are the speeds of its constituent particles (see Box 2.2). This is the main reason why nuclear reactions can only take place at temperatures exceeding a million Kelvin, which are only present at the centres of stars.

Box 2.7 Find a Hotter Place

In 1920, the famous British astronomer Arthur Eddington studied the equations which describe the structure of stars like the Sun, and determined that the temperatures at the centres of stars are higher than 10 million K. He suggested that these high temperatures along with the high density of the gas there, are just right for nuclear reactions to occur and these reactions, in particular those which convert hydrogen to helium, should be the source of energy of these stars. Eddington's ideas were not accepted by the physicists at that time who could not believe that hydrogen nuclei (protons) can come close enough to undergo nuclear reactions because of the electrostatic force of repulsion among them. Eddington however firmly believed that with the high velocities that protons would have at the temperature of more than 10 million K, there is a good chance for them to come sufficiently close in spite of their electrostatic repulsion. He confidently wrote in one of his books that "The helium that we handle must have been put together at some time and some place. We do not argue with the critic who urges that the stars are not hot enough for this process; we tell him to go and find a hotter place."

The effective hydrogen burning reaction can be described as

$$4\,\mathrm{p} \rightarrow \alpha + 2\,\mathrm{e}^{+} + 2\,\nu_{\mathrm{e}} + 2\,\gamma$$

Here, p, α, e^{+}, ν_e and γ stand for proton, alpha particle, positron, electron neutrino, and gamma ray photon, respectively. A positron is an elementary particle (see Box 2.8), which has the same mass as that of an electron but has a unit positive charge instead of the unit negative charge that the electron has. It is also called the antiparticle of the electron. An electron neutrino is an elementary particle having no charge and almost no mass. A gamma ray photon is a particle equivalent of high energy, gamma radiation. The alpha particle, the nucleus of helium having two protons and two neutrons is also denoted by ^{4}He. The superscript 4 indicates mass number, i.e., the total number of nucleons in the helium nucleus and He is the symbol for the element helium. In the above reaction, 4 protons are the reactants and alpha particle, positrons, and electron neutrinos are the products. The total mass of the reactants is larger than the total mass of the products and the difference is emitted as energy in the form of gamma rays. Note that two protons have been converted to two neutrons in the process of the nuclear reaction and two positive charges have been emitted in the form of positrons so that the total charge is conserved. An electron neutrino is always emitted whenever a proton gets converted to a neutron. The reaction is pictorially shown in Fig. 2.4.

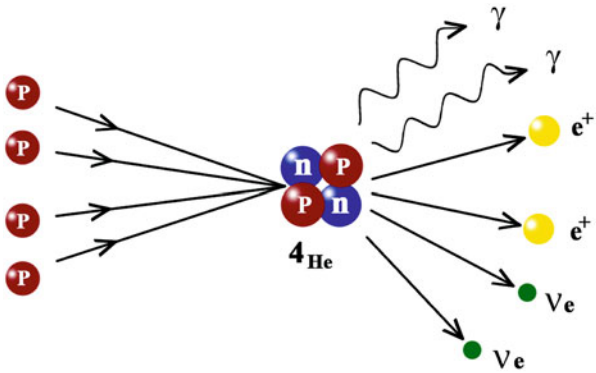


Fig. 2.4 Figure showing the effective hydrogen burning reaction: four protons combine to form helium nucleus (alpha particle). Two protons get converted to neutrons in the process and two positrons, two electron neutrinos, and two gamma rays are emitted

Box 2.8 Elementary or Fundamental Particles

What are the smallest particles which make up all the matter around us? Dalton was first to propose in 1803 that all matter is made up of atoms; atom meaning indivisible in Greek. In 1897, Thompson discovered the electron and in 1917 Rutherford discovered the proton, giving rise to the theory that every atom has a small, positively charged nucleus containing protons and is surrounded by electrons. Neutrons were discovered by Chadwick in 1932. For a long time, protons, neutrons and electrons were thought to be the elementary particles from which all matter is built. Soon other particles were discovered. The positron, which is an elementary particle having the same mass as that of an electron but an opposite charge, was discovered in 1932. The neutrino, which is also an elementary particle having very small mass and no charge, was theoretically proposed by Pauli in 1930 and was experimentally discovered in 1956. These added to the list of elementary particles.

Unlike these particles, the presence of some other elementary particles was proposed purely on the basis of theoretical considerations. These particles were never discovered experimentally. According to this theory, though electrons are indeed elementary particles, neutrons and protons are not. They are composite particles made up of smaller particles called quarks which are believed to be elementary particles. Theoretical models based on the presence of quarks explain the properties and interactions of protons, neutrons and other particles known to exist in nature. There are believed to be 6 types of quarks which have been named up, down, top, bottom, strange and charm. These come in three “colours” each, namely red, blue, and green. Note that the “colour” just describes one particular property of the quarks and has nothing to do with the colours that we see with our eyes. One interesting property of quarks is that they have fractional

(continued)

Box 2.8 (continued)
charges. Their charges have values $-1/3$ and $+2/3$ while protons, electrons and positrons have unit charges. Another interesting property of quarks is that they can never be free particles and are doomed to be bound inside protons, neutrons and other particles under ordinary circumstances. This is because, the attractive (strong) force between two quarks grows with distance between them. Thus, the larger the distance, the higher is the attractive force making it impossible for them to go away from each other and become free. One exception to this occurred in the very early Universe when the temperatures and densities were so high that the quarks could exist as free particles. We will learn about this in Chap. 8.

In addition to the six quarks and their antiparticles, i.e., particles having identical masses but opposite charge and other properties, there are six elementary particles collectively called leptons. These are electron, muon, tau, and the corresponding neutrinos, i.e., electron neutrino, muon neutrino, and tau neutrino. These 6 leptons also have their antiparticles. In addition, there are 6 types of yet another type of particles called bosons. These are gluon, photon, graviton, Higgs, $W^{\pm}$, and Z. The gluons, like quarks, can not exist as free particles and are always confined inside protons, neutrons, and other similar particles except in the very early Universe.

Thus, the total number of elementary particles from which all the matter and radiation forms is 18 (and their antiparticles). Protons are for example formed from 3 quarks having charges +2/3, +2/3, and $-1/3$ with a total charge of +1, while neutrons are made up of quarks having charges +2/3, $-1/3$, and $-1/3$, giving zero total charge to them.

2.2.2 Award Winning Work: Energy Generation in Stars

That nuclear reactions are a possible source of energy inside stars was proposed by Arthur Eddington in the 1920s. However, it was not accepted by scientists for the reason that it is extremely difficult for protons to come close enough for the hydrogen burning nuclear reaction to take place. Bethe was a theoretical nuclear physicist. He extensively studied a large number of nuclear reactions which could possibly take place in stellar interiors. He meticulously determined the rates at which these reactions would take place, the particles and the amount of energy which would be emitted during each reaction, and took into account the amounts of various elements present in the stellar material. He showed that conversion of hydrogen to helium can take place in two different ways. Each way involved a chain of nuclear reactions. The two chains are called PP chain and CNO cycle. Both chains effectively convert 4 protons into helium nucleus, releasing energy in the form of high energy gamma rays, and elementary particles in the process as described above.

The PP chain starts with the reaction of two protons to produce a deuterium nucleus according to the reaction

$$2\,\mathrm{p} \rightarrow {}^{2}\mathrm{D} + \mathrm{e}^{+} + \nu_{\mathrm{e}} + \gamma.$$

Here, ^{2}D is a deuterium nucleus having mass number of 2. It has one proton and one neutron and is an isotope of hydrogen. In the reaction, one proton has been converted to a neutron with the emission of a positron, an electron neutrino and a gamma ray photon. So far, two protons have been used.

This is followed by four further nuclear reactions. In the first, the deuterium nucleus reacts with a proton to produce the nucleus of ^{3}He according to

$$^{2}\mathrm{D} + \mathrm{p} \rightarrow {}^{3}\mathrm{He} + \gamma.$$

^{3}He is an isotope of ^{4}He and has two protons and one neutron. Note that a third proton has been used in this reaction. No conversion of proton to neutron has taken place and only a gamma ray is emitted.

In the next reaction, the ^{3}He nucleus reacts with one of the already existing ^{4}He nucleus to generate a nucleus of beryllium, ^{7}Be according to

$$^{3}\mathrm{He} + {}^{4}\mathrm{He} \rightarrow {}^{7}\mathrm{Be} + \gamma.$$

The beryllium nucleus has four protons and three neutrons. No conversion among nucleons takes place in this reaction. The next reaction is between the beryllium nucleus and an electron, producing a lithium nucleus, ^{7}Li as

$$^{7}\mathrm{Be} + e^{-} \rightarrow {}^{7}\mathrm{Li} + \nu_{e}.$$

^{7}Li has three protons and four neutrons. A proton has been converted into a neutron in this reaction and a neutrino is emitted.

The last step in the PP chain is the reaction of the lithium nucleus with a proton, resulting in two alpha particles according to

$$^{7}\mathrm{Li} + {}^{1}\mathrm{H} \rightarrow 2\,{}^{4}\mathrm{He}.$$

No conversion among nucleons takes place in this reaction. A fourth proton has been used in this reaction and the total of four protons have given us an alpha particle. The existing alpha particle used in the third reaction in the chain is returned back to the stellar material. If we look at the five reactions carefully we see that effectively we have converted four protons and one electron into an

alpha particle with the emission of two neutrinos, one positron and energy in the form of gamma rays in a way very similar to the effective hydrogen burning reaction described above and shown in Fig. 2.4. Bethe had also proposed a somewhat different last step but the above is the accepted reaction now.

Bethe's calculations showed that the PP chain can produce sufficient energy to explain the thermal equilibrium and the luminosity of stars having masses lower than the mass of the Sun. He also found that the PP chain would not be able to provide the amount of energy required for the thermal equilibrium of heavier stars. He realized that heavier nuclei have to be involved in the nuclear reactions occurring inside these stars. After considering the abundances, i.e., the relative quantities of various nuclei in stellar material and rates of reactions involving these elements, he established a chain of 6 reactions which can take place in heavier stars and provide sufficient energy for the thermal equilibrium of these stars. These reactions start by adding one proton to an already existing carbon nucleus and then adding three more protons one by one to the products, producing nitrogen and oxygen along the way and finally producing ^{4}He and returning the initial carbon nucleus back to the gas. Thus, carbon acts merely as a catalyst in this chain. These reactions are as follows.

$$^{12}\mathrm{C} + \mathrm{p} \rightarrow {}^{13}\mathrm{N} + \gamma$$

$$^{13}\mathrm{N} \rightarrow {}^{13}\mathrm{C} + \mathrm{e}^{+} + \nu_{\mathrm{e}}$$

$$^{13}\mathrm{C} + \mathrm{p} \rightarrow {}^{14}\mathrm{N} + \gamma$$

$$^{14}\mathrm{N} + \mathrm{p} \rightarrow {}^{15}\mathrm{O} + \gamma$$

$$^{15}\mathrm{O} \rightarrow {}^{15}\mathrm{N} + \mathrm{e}^{+} + \nu_{\mathrm{e}}$$

$$^{15}\mathrm{N} + \mathrm{p} \rightarrow {}^{12}\mathrm{C} + {}^{4}\mathrm{He}$$

In these reactions, C, N, and O are the symbols for carbon, nitrogen, and oxygen respectively. They have 6, 7, and 8 protons respectively in their nuclei. The first reaction adds a proton to an existing carbon nucleus, which subsequently gets converted to nitrogen and oxygen nuclei and is finally returned back as carbon. In the first, third, fourth, and sixth reactions each, a proton is added to the other reactant. No conversion among nucleons takes place in these reactions and only energy is emitted. In the second and fifth reactions a proton in the nucleus of nitrogen and oxygen each is converted to neutron. This chain of 6 reactions is called the CNO cycle. Like the PP chain,

the CNO cycle also effectively converts 4 protons into a helium nucleus, in the process, releasing particles and energy in the form of high energy radiation. Bethe showed that the energy generated in the CNO cycle is sufficient to explain the luminosity of stars heavier than the Sun.

In addition, Bethe also showed that elements heavier than helium can not be produced inside stars. For producing heavier elements, protons would have to be added to helium nucleus one by one. He showed that all elements up to boron disintegrate on reacting with a proton. He also considered the so called triple alpha reaction which can generate a carbon nucleus from 3 alpha particles as

$$3\,{}^{4}\mathrm{He} \rightarrow {}^{12}\mathrm{C} + \gamma$$

He showed its probability to be extremely small for the reaction to be of any consequence. We will see in the next chapter how this problem of generating elements heavier than helium was solved. Bethe published his results in 1938–1939. His work solved the decades old puzzle of the source of energy generation in stars and he was awarded the Nobel Prize of 1967 for this work.

Subsequently astronomers have established two more branches to the PP chain. The three branches are called PPI, PPII and PPIII chains, with Bethe's chain being the PPII chain. It was also shown that most of the energy generation in the Sun and stars of similar masses occurs through the PP chain, while the CNO cycle is important for stars with masses larger than about 1.3 $M_{\odot}$.

2.3 Nobel Prize 2002: R. Davis Jr. and M. Koshiba

The Nobel Prize for the year 2002 was given to Raymond Davis Jr. and Masatoshi Koshiba for "Pioneering contributions to astrophysics, in particular for the detection of cosmic neutrinos." They shared half of the prize amount; the other half was awarded to Riccardo Giacconi that year.

Credit: Davis: By National Science Foundation—[1] [2], Public Domain, https://commons.wikimedia.org/w/index.php?curid=20024888

Raymond Davis Jr. was born in the USA in 1914. He graduated from the University of Maryland in chemistry and later completed his Masters and Ph.D. from Yale University in 1942. After spending the next three years in the US army, he worked at the Monsanto's Mound Laboratory, in Ohio. In 1948 he joined the Brookhaven National Laboratory.

At Brookhaven, Davis initially worked on detecting the neutrino flux from nuclear reactors using the same technique that he later used for the detection of neutrinos coming from the Sun. Later, he was involved in measuring the argon content of rocks on the Earth, of meteorites, and later of lunar rocks. After retiring from Brookhaven in 1984, he was a research professor at the University of Pennsylvania. He passed away in 2006.

Credit: Koshiba: By 首相官邸ホ ム ペ ジ, CC BY 4.0, https://commons.wikimedia.org/w/index.php?curid=52375094

Masatoshi Koshiba was born in Japan in 1926. During his high school days, he came across a book which contained a dialogue with Albert Einstein. This attracted him towards science. He graduated from the University of Tokyo in 1951 and obtained his Ph.D. in physics from the University of Rochester,

USA in 1955. He returned to Japan in 1958 after a stint at the University of Chicago. In Japan, he was a professor at the University of Tokyo, where he remained till 1987. During this period he held directorship of several institutes including some in the US. After retiring in 1987, he remained a professor at Tokai University till 1997. He passed away in 2020.

2.3.1 Background

Calculations of stellar structure had started in the late nineteenth century, but picked up pace in the twentieth century after the source of energy generation in the stars was understood. These calculations are very complex as they involve a large number of processes. One has to consider all possible nuclear reactions that can take place in the star (as discussed in Sect. 2.2.2), the rate at which they generate energy, the transport of the generated energy from the centre to the surface of the stars, the balance between the gas pressure and gravity, the relation between pressure, temperature and density, the state of ionization of the gas (i.e., how many electrons are removed from different types of atoms) at a given temperature and density etc.,– all these at all points inside the star. The availably of computers made detailed stellar structure studies possible.

The starting parameters for such calculations for a star undergoing hydrogen burning at its centre, are the mass of the star, its chemical composition, and its age. In this section we will ignore the age factor and assume that the chemical composition of the star is uniform throughout its interior and is same as that at the surface which can be measured observationally and is its initial composition. The age of the star is of course important and will be discussed in the next chapter. One then has to solve the four basic equations of stellar structure at every point inside the star. These describe (i) the hydrostatic equilibrium (see Sect. 2.2.1), (ii) the mass distribution, i.e., how the particle density changes with distance from the centre, (iii) the nuclear energy generation rate, and (iv) energy transport from centre to the surface. All the necessary atomic and nuclear physics details have to be provided. The stellar model calculations then return the values of temperature and density throughout the star as well as the star's observable properties like the surface temperature, spectrum, size, and luminosity of the star. These can be compared with their observed values. If the mass of the star is not known, its value used in the model calculations can be varied till the model predictions are in agreement with the observed properties of the star. Such stellar models, thus, enable us to determine the mass of a star and understand the physical conditions in the interiors of the star.

From such stellar models we have understood, as noted in Sect. 2.2.2, that for low mass stars with masses up to 1.3 $M_\odot$, i.e., 1.3 times the solar mass, the temperatures and densities at the centre are such that the PP chain plays the dominant role in the generation of energy. Thus, the energy at the centre of the Sun is mostly produced by the PP chain. On the other hand, the conditions at the centres of heavier stars are conducive for the CNO cycle to be the dominant mode of energy generation.

Solar Neutrinos

From the solar model, we know that PP chain is responsible for the generation of energy at the centre of the Sun. This is an indirect conclusion drawn on the basis of the agreement between the observed properties of the Sun and the predictions of the solar models. The model tells us that the temperature at the centre of the Sun is about 2.7×10^7 K while on its surface, the temperature has been measured to be about 5800 K. There is a gradual decrease in temperature as we go from the centre to the surface. The energy is generated by nuclear reactions which can only take place in the central, core region because the necessary high temperatures and densities are present only in that region. The energy generated there is in the form of high energy gamma ray photons. These photons can not directly come out of the Sun. The material in the interior of the Sun is dense and the gamma rays get absorbed by the material after traveling a small distance away from the centre. They deposit their energy in the material there. That material in turn emits radiation, but of somewhat smaller average energy as the temperature of the material there is lower than that at the centre (see Box 2.3). These rays again get absorbed on the way out only to be reemitted at a still lower average energy corresponding to the lower temperature of the solar material there. This goes on until the radiation reaches the surface layer of the Sun called the photosphere. Photosphere means "sphere of light" and is the sphere that we see when we look at the Sun. The solar material in the layer on the outside of the photosphere is not sufficiently dense and the radiation emitted by the material of the photosphere, which is the black body radiation corresponding to the temperature of the photosphere, can escape the Sun. This is the radiation that reaches us. Thus, all information that we can get by analyzing the solar radiation that reaches us is about the photosphere or the surface of the Sun and about the gases outside it. The cooler material which exists outside the photosphere is a part of the atmosphere of the Sun. The layer immediately outside the photosphere is called the chromosphere, and the layer beyond the chromosphere is called

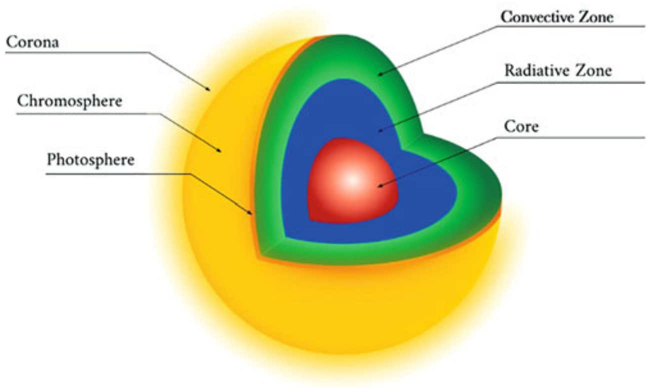


Fig. 2.5 Some regions of the Sun. The central core where energy generation takes place is shown in pink. Also shown are the photosphere, chromosphere, and corona. The convective zone and radiative zone are regions in which the transfer of heat occurs through convection and radiation, respectively

the corona. These layers are schematically shown in Fig. 2.5. The temperature keeps decreasing from the centre to some point in the chromosphere where it reaches close to 4300 K, and then increases again, up to 3 million K in the corona.

Is there a way to look directly at the centre of the Sun so that we can get first-hand information about the conditions there? The answer, surprisingly, is yes. Though the radiation emitted at the centre can not reach us directly because of its absorption by the solar material, there exists a type of particle which is emitted during the nuclear reactions and which can come out of the Sun without getting absorbed, and reach the Earth. This is the electron neutrino. We have seen that these neutrinos are emitted in the nuclear reactions in the PP chain (see Sect. 2.2.2). They are essentially emitted when protons are converted to neutrons. As the other PP chains also carry out such a conversion, neutrinos are emitted in these chains as well. These neutrinos, emitted in the core of the Sun, can reach us unabsorbed. Why can the neutrinos come out directly without getting absorbed, while the radiation (photons) can not? The reason is that while the radiation interacts with matter through electromagnetic force, the neutrinos experience only weak force. At the neutrino energies of interest, i.e., at the energies of the neutrinos emitted in the PP chain, the weak force is about 10^{11} times weaker than the electromagnetic force. The interaction of neutrinos with matter is thus, much weaker than that of photons and the probability of neutrinos getting absorbed by the solar material is 10^{22} times smaller (as the probability depends on the square of the strength of the force

of interaction) than that of photons. This is the reason why neutrinos do not get absorbed by the solar material and can come out of the Sun directly. Thus, if we can measure the flux of the solar neutrinos reaching the Earth, i.e., the number of neutrinos received per unit area per unit time, we can get direct information about the rates of nuclear reactions taking place at the centre and verify their consistency with the results of the solar model.

2.3.2 Award Winning Work: Detection of Cosmic Neutrinos

The Sun produces about 10^{30} neutrinos every second. Out of these, a fraction falls on the Earth. Approximately 100 billion neutrinos pass through an average finger nail on the Earth every second. They pass through our bodies and even the Earth, mostly unhindered, our bodies absorb about one neutrino in a lifetime. How then, does one detect these neutrinos? For detecting them, they have to interact and be absorbed by the material of the detector. But as their interaction with matter is extremely weak, they are also extremely unlikely to get absorbed and get detected. It required big thinking to devise an experiment to detect them. This is exactly what Raymond Davis and John Bahcall did when they proposed an experiment for the purpose in 1964. The idea was to make a detector having a huge quantity of neutrino absorbing material. As solar neutrinos pass through this huge amount of material, at least some of them would get absorbed. Counting the number of absorbed neutrinos and knowing the probability of such absorption, one can theoretically work out how many neutrinos had passed through the detector material and therefore determine the number of neutrinos emitted per second by the Sun.

Davis set up the experiment, known as the Homestake experiment, in the Homestake Gold Mine in the USA. The mine was 1,478 m deep. A deep mine is necessary for the successful detection of the solar neutrinos, as explained in Box 2.9. A huge tank having a capacity of 380 m^3, i.e., about 102,000 gallons, was constructed and filled with perchloroethylene which is a common dry-cleaning fluid. Neutrinos from the Sun can interact with chlorine nuclei, which have 17 protons and convert them into argon nuclei having 18 protons, converting one neutron into proton and emitting an electron in the process.

Box 2.9 Cosmic Rays and Neutrino Detector

Cosmic rays are charged particles which are emitted by cosmic sources and arrive at the Earth. These particles have very high energies and travel with speeds which are comparable to the speed of light. Until a few years back they were the only entities coming from extraterrestrial sources which we could observe, in addition to the electromagnetic waves and neutrinos from the Sun and from supernova 1987 about which we will learn in the next chapter. The fourth such entity which was first detected in 2015 consists of the gravitational waves, about which we will learn in Chap. 7.

The charged particles in cosmic rays are mainly protons (about 90%) and alpha particles, i.e., helium nuclei (9%), the rest (1%) being nuclei of heavier elements. They were discovered by Victor Hess in 1912. The cosmic rays coming from extraterrestrial sources which include the Sun, supernovae, the Milky Way, and possibly other galaxies are called primary cosmic rays. When these high energy particles enter the Earth's atmosphere, they interact with atmospheric gases and produce a cascade of particles which continue to move forward and rain down on the Earth. These are called secondary cosmic rays and such events are called air showers. The number of cosmic ray particles hitting each square metre every second on the surface of the Earth is in the tens of thousands. The cosmic ray showers contain electron neutrinos and muon neutrinos which are elementary particles and are similar in properties to electron neutrinos (see Box 2.8). The ratio of the number of muon neutrinos to that of electron neutrinos is about 2:1. These neutrinos are known as atmospheric neutrinos.

If a solar neutrino detector was placed on the surface of the Earth, the cosmic rays would interact with the detector material and produce a strong signal which will completely drown any signal coming from solar neutrinos. This is the reason why the solar neutrino detector had to be set up in a deep mine. Most cosmic rays can not penetrate too deep inside the Earth. Thus, having the neutrino detector underground at a large depth below the surface of the Earth helps in cutting down the flux of cosmic rays and increases the chances of detecting solar neutrinos.

The argon nuclei remain intact on the average for 35 days, after which they undergo radioactive decay and convert back to chlorine by emitting a positron and a neutrino. Argon is a noble gas and it is easy to separate chemically from the chlorine-rich solvent. In the experiment, the argon atoms were separated and counted every few weeks. This gave the number of solar neutrinos which reacted with chlorine in the tank during the period between two separations, and one could calculate the number of reactions per second. Knowing the probability of absorption of neutrinos, this number could be converted into the number of neutrinos which passed through the tank and thereby into the number of neutrinos emitted by the Sun per second. This number could be compared with the number predicted by the solar models and the validity

Fig. 2.6 A view of the Homestake experiment. **Credit:** https://commons.wikimedia.org/wiki/File:HD.6D.653_(12000067315).jpg ENERGY.GOV, Public domain, via Wikimedia Commons

of these models could be checked. A picture of the experiment is shown in Fig. 2.6.

The first results of the experiment were published in 1968 and were very puzzling. The solar neutrino flux measured by Davis was significantly lower than the flux expected on the basis of solar models. This gave rise to what was called the solar neutrino puzzle (see Box 2.10).

Box 2.10 Solar Neutrino Puzzle

The results of the Homestake experiment were eagerly awaited by astronomers, especially the solar physicists. But when the results were announced in 1968, they were very puzzling. The neutrino flux was significantly lower than the flux predicted by the solar models.

The results seemed to suggest that the solar models were wrong and the values of temperature and density in the core of the Sun as suggested by the models were too high. This meant that the actual gas pressure at the centre must be lower than what was predicted by the models, and some additional pressure would be necessary to balance the force of gravity. A lot of activity ensued among the theoreticians trying to find ways of reducing the theoretically predicted neutrino flux while keeping all other properties of the Sun unchanged as they matched well with the observations. Later astronomers developed another method to measure the temperature and density inside the Sun through

(continued)

Box 2.10 (continued)

observations of seismic waves on the surface of the Sun. These indicated that the central temperature and density in the Sun were consistent with values given by the solar models.

What could then be happening to the neutrinos? Why were we not detecting them in the expected numbers? Could they be getting absorbed inside the Sun? That was highly unlikely due to the weakness of the force with which they interact with matter. The answer lay in the physics of elementary particles. Neutrinos come in three different types. These are electron neutrinos, muon neutrinos, and tau neutrinos, muon and tau being elementary particles of the same class as electrons (see Box 2.8). The neutrinos produced in the Sun are electron neutrinos and the Homestake experiment and Kamiokande (described below) were designed to detect this type of neutrinos. In 1968, it was suggested by theoreticians that the neutrinos could change their type with time, i.e., get converted into neutrinos of other types. Thus, the electron neutrinos can get converted into muon or tau neutrinos and vice versa the muon and tau neutrinos can get converted into electron neutrinos. This is known as neutrino oscillations. As the experiments only detect the electron neutrinos, they may be missing those which got converted into other types of neutrinos on the way to the Earth. However, the conversion rate was found to be too small to explain the discrepancy between the observed and predicted neutrino flux.

In late 1970s and 1980s, it was shown that the passage of neutrinos through solar material amplifies neutrino oscillations and the effect of oscillations could be large enough to explain the shortage of detected electron neutrinos. This effect of matter on neutrino oscillation is known as the MSW effect after the physicists who discovered it, namely, Mikheyev, Smirnov, and Wolfenstein. An experiment to detect neutrinos of all types coming from the Sun was set up in Creighton mines in Sudbury, Canada. This used heavy water (which is water having deuterium in place of hydrogen) which can react with and detect all three types of neutrinos. The first results of the experiment were published in 2001. The total observed number of all three types of neutrinos matched well with the number of neutrinos emitted by the Sun as predicted by the solar models. This gave firm evidence of neutrino oscillations. About 2/3rds of the electron neutrinos get converted into neutrinos of other types while traveling from the centre of the Sun to the Earth and could not be detected by the Homestake experiment and Kamiokande. Thus, the origin of the solar neutrino puzzle lay in a lack of understanding of neutrino properties and not in incorrect solar models. The experimental confirmation of neutrino oscillations also proved that at least one of the three types of neutrinos must have mass as otherwise neutrino oscillations can not take place. That neutrinos have mass was earlier suggested by the difference in arrival times of different types of neutrinos from the 1987 supernova explosion (see Box 3.3).

Koshiba was a cosmic ray physicist. In 1969 he set up an underground experiment in Kamioka mines in Japan, to detect cosmic rays. In 1978 he set up another experiment in the same region for detecting the radioactive

decay of protons which was predicted by some theories of elementary particle physics. For this experiment, he had to dig another mine and put a tank containing 30 tons of pure water at a depth of 1 km below the surface of the Earth. The experiment, which was called Kamioka Nuclear Decay Experiment (Kamiokande in short), started taking data from 1983 but failed to detect any proton decay. Koshiba then modified his experiment to detect solar neutrinos.

Koshiba's detector used a technique which was different from that used by Davis. His detector had a huge stainless steel tank which had a diameter of 39 m. This was filled with 50,000 tonnes of purified water. Neutrinos passing through such a tank occasionally interact with the hydrogen in the water molecules, releasing electrons. These electrons have speeds which are larger than the speed of light in water. Remember that according to the special theory of relativity, nothing can travel faster than the speed of light in vacuum. The speed of light in a material medium is smaller than its speed in vacuum. For example, in water it is about 3/4ths of its speed in vacuum. The speeds of electrons produced by the neutrinos passing through water are higher than this and they emit a flash of light known as Cerenkov radiation. More than 13,000 sensitive light detectors surrounding the tank (see Fig. 2.7) watch for these flashes. They can pinpoint whether the neutrino which caused the detected Cerenkov radiation originated from the Sun, or came from other directions, and could identify atmospheric neutrinos which come from all directions. Kamiokande started taking data from 1986. However, before it could detect solar neutrinos, it detected 12 neutrinos in a short interval of time of about 13 s, in 1987. These neutrinos were not coming from the Sun. Later they

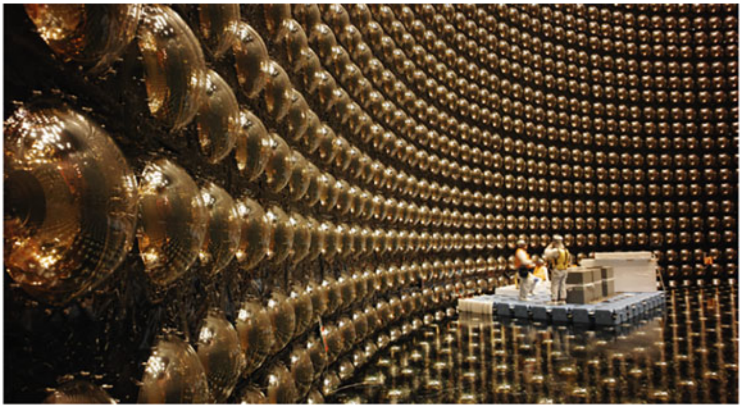


Fig. 2.7 A view of the Kamiokande neutrino experiment. **Credit:** Kamioka Observatory, ICRR, University of Tokyo

were confirmed to have come from a supernova explosion in the nearby galaxy, the Large Magellanic Cloud. The observation of these neutrinos gave support to the theory of supernova explosions. This was the first and is so far the only cosmic source of neutrinos detected outside the solar system. Koshiba retired after that but continued to work as consultant and collaborator. The collaboration finally announced the detection of solar neutrinos in 1989 and confirmed Davis's results. Later, the Kamiokande detector was upgraded and is now called super-Kamiokande.

The super-Kamiokande detector could not only detect atmospheric neutrinos, but could also distinguish between the electron neutrino and the muon neutrino. The ratio of the two types of atmospheric neutrinos, the muon and electron neutrinos, was determined and found to be much smaller than the expected value of 2 (see Box 2.9). This so called atmospheric neutrino anomaly was explained as being due to neutrino oscillations.

Both the Homestake and Kamiokande were difficult to conceive and to set up. The results of these experiments provided crucial support to the solar models and also to the theory of supernova explosions. In addition, these results led to a new understanding of the properties of neutrinos. Koshiba was awarded the Nobel Prize for detection of cosmic neutrinos along with Davis in 2002.

3

Stellar Evolution

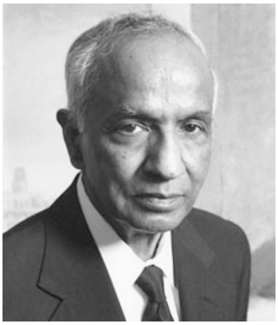
3.1 Introduction

The Nobel Prize in the category of stellar evolution was given in the year 1983 to two astronomers: Subrahmanyan Chandrasekhar and William Alfred Fowler. The work of Chandrasekhar is purely theoretical and is in the field of structure of white dwarf stars, which are one of the end stages in the life of stars. This work was done in the early nineteen thirties when Chandrasekhar was a fresh graduate of physics. The work by Fowler covers both experimental and theoretical nuclear physics as well as astrophysics, and involves the synthesis of elements heavier than helium during the evolution of stars. The work was spread over several decades, a major portion having been performed before the nineteen sixties.

3.2 Nobel Prize 1983: S. Chandrasekhar

The Nobel Prize for the year 1983 was given to Subrahmanyan Chandrasekhar "for his theoretical studies of the physical processes of importance to the structure and evolution of the stars." He shared the prize with W. A. Fowler.

P. Khare, *Nobel Prizes in Astronomy*, Astronomers' Universe,
https://doi.org/10.1007/978-3-031-29639-0_3



Credit: Hanna Holborn Gray Special Collections Research Center, University of Chicago Library.

Subrahmanyan Chandrasekhar was born in Lahore in 1910. His mother is credited for arousing his curiosity at an early age. Nobel Prize winner C. V. Raman was Chandrasekhar's uncle. Chandrasekhar was home schooled till the age of twelve. Later he did his graduation from the Presidency College, Chennai. He wrote his first research paper during that time. In 1930 he was awarded a government fellowship to pursue doctoral studies at Cambridge. He obtained his doctorate in 1933.

Chandrasekhar left England for the USA in 1936. Initially he worked in the Yerkes Observatory. Later he shifted to the University of Chicago. There he became a full professor at the age of 33. He continued his research there till his death in 1995. He was the Editor of the Astrophysical Journal, arguably the most prestigious journal in the subject, from 1952 to 1971.

Chandrasekhar worked in several areas of astrophysics. He used to study a topic in depth, publish research papers on it, and finally write a masterly book on the subject before moving on to a new topic of study. He has authored ten scholarly books and several lecture notes. He rewrote Newton's Principia in a way which would be understandable to a general reader. He was also fond of literature and arts along with science, and has written a book titled "Truth and Beauty: Aesthetics and Motivations in Science".

3.2.1 Background

Hydrogen burning reactions are taking place at the centre of the Sun since 4.5 billion years, and theoretical calculations tell us that they will continue for a similar duration in the future. This is the source of energy for the Sun. Hydrogen burning is the first source of nuclear energy for all stars as (i) hydrogen is the most abundant element in all stars and (ii) hydrogen burning

occurs at the lowest temperatures; burning of heavier elements require higher temperatures. This is because nuclei of elements heavier than hydrogen possess larger numbers of protons and hence larger positive charge. For two such nuclei to come close enough (10^{-13} cm) to be able to undergo fusion, they will have to overcome a much higher electrostatic force of repulsion (remember, the electric force is proportional to the product of the charges of the two interacting particles) and so will need very large speeds, i.e., very high temperatures (see Box 2.2). All stars pass through the hydrogen burning phase. The longest period in the life of a star is spent in this phase.

Obviously, this phase can not last for ever. The amount of hydrogen in the core goes on decreasing with time. This results in some changes in the structure of the star. This is why, as was mentioned in the last chapter (see Sect. 2.3.1), the age of a star enters into its structure calculation. There, the age of the star was ignored for ease of understanding. More accurate calculations of stellar models do take into account the effect of change in the chemical composition in the stellar core with age of the star. Eventually, all the hydrogen at the centre of a star gets exhausted, i.e., gets converted into helium. What happens after that? The hydrogen burning nuclear reactions stop for want of hydrogen, i.e., for want of a nuclear fuel, and the energy generation in the star ceases. The star being at a very high temperature continues to lose energy to the cooler surroundings, i.e., it continues to radiate away its thermal energy and so starts cooling down. The moment this happens, the gas pressure starts falling as it is proportional to temperature for all gases found in the Universe (except under special circumstances in the interiors of some stars, as we will soon find out). The balance between the gas pressure and the force of gravity can not be maintained, gravity takes over, and the star starts contracting.

As discussed earlier (see Sect. 2.2.1), when a star contracts, its gravitational potential energy, which depends on the mass and radius of the star, decreases. This decrease results in an increase in thermal energy of the gas and the star begins to heat up again. This contraction and increase in temperature continues until the temperature becomes high enough for the next nuclear fuel to start burning.

After hydrogen, the next lightest and most abundant element is helium. Helium can undergo fusion through the triple alpha reaction (see Sect. 2.2.2), three helium nuclei, i.e., alpha particles, fusing to form a carbon nucleus. Three alpha particles together have 6 protons and 6 neutrons, the same as a carbon nucleus. Thus, there is no conversion of protons into neutrons or vice versa, and only energy in the form of gamma rays is released in the process as the mass of 3 alpha particles together is larger than the mass of a carbon nucleus. Clearly, for even two alpha particles to come close enough to undergo

nuclear reaction, they have to overcome 4 times as much repulsive force as that which has to be overcome by two hydrogen nuclei (protons) to undergo nuclear reactions. Thus, helium fusion requires much higher temperatures, about 2×10^8 K, than that required for hydrogen burning which is 4×10^6 K (for PP chain). Bethe had shown the probability of this reaction to be very small. We will learn how this reaction can indeed take place inside stars, in Sect. 3.3.2. Some of the carbon nuclei thus formed react with an alpha particle to generate oxygen as

$$^{12}\mathrm{C} + {}^{4}\mathrm{He} \rightarrow {}^{16}\mathrm{O} + \gamma$$

Helium burning, which generates carbon and oxygen, is the second cycle of nuclear fusion, the first being hydrogen burning. After the helium in the centre gets exhausted, the same processes that took place after the first cycle, i.e., contraction and heating of the star, take place till the central temperature becomes high enough for the next cycle of nuclear reactions to begin. The third cycle converts carbon into oxygen, sodium, magnesium and neon. There can be several such cycles. Each new cycle involves nuclei with higher numbers of protons, and hence requires higher temperatures.

These cycles can, in principle, continue till all material in the core of the star is converted into iron. Till then, each fusion reaction results in products which have a total mass which is smaller than that of the reactants and the mass difference is released as energy. This does not happen for the fusion of iron or heavier elements due to the fact that the binding energy per nucleon is maximum for iron (see Box 2.6). So these reactions can not be sources of energy for the stars.

All fusion cycles till the generation of iron do not take place in all stars. How many fusion cycles a star undergoes depends on its initial mass. Stars with masses smaller than about 8 $M_\odot$ undergo up to three cycles, the number of cycles increasing from one for stars with lowest masses to three for stars with masses close to 8 $M_\odot$. The reason why no further cycle can take place is that the star never achieves the necessary high temperature for the next cycle to start. After the nuclear fuel is exhausted and energy generation stops, the star, as seen above, starts contracting. How can the contraction ever stop in absence of any further nuclear energy generation?

As we know, at the high temperatures and densities in the cores of these stars all atoms are fully ionized, i.e., they are completely stripped of their electrons. This in fact allows the densities in the cores of these stars to reach very high values; the atoms have much bigger sizes and can not be squeezed close to one another beyond a limit, while nuclei are much smaller in size

and it is possible to squeeze them together and achieve higher densities. The electrons are completely free and form what is called free electron gas. When the densities in the cores become higher than a certain limiting value, the laws of classical physics do not hold good for the free electron gas. New physical laws which come under the branch of physics called quantum mechanics come into play. Under these new laws, the free electron gas changes its behaviour. It loses the properties typical of a classical gas and is called degenerate gas (see Box 3.1), meaning gas which exhibits properties which are not typical of its kind. The pressure of such a gas does increase with increase in particle density, but it does not depend on its temperature. After this stage is reached, the electron gas pressure keeps increasing as the star contracts, due to the increase in electron density. If at some point during the contraction pressure can balance gravity, then the contraction stops. The balance is then maintained for ever, as even though the star keeps radiating its thermal energy and its temperature keeps decreasing in absence of any nuclear energy generation, its pressure does not change as the pressure no longer depends on temperature. The degenerate electron gas pressure would keep balancing gravity for ever. Such a star, whose gravity is balanced by the pressure of degenerate electron gas, can be considered a dead star as its composition and size remain unchanged for ever. This is the end stage in the life of all stars which have initial masses smaller than about 8 $M_\odot$.

Box 3.1 Degenerate Gas

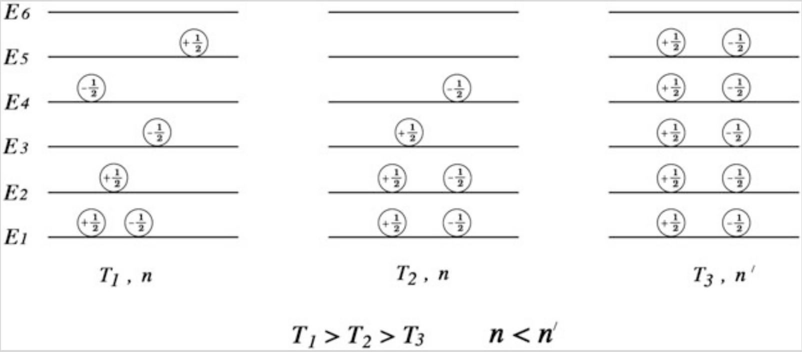
According to quantum mechanics the energy of a particle which is confined within a certain volume can only have certain discrete values. A particle having one of these discrete values is said to be in an energy state or energy level having that particular value of energy. All elementary particles have a property called spin. For electrons the spin is 1/2 (in some units). A particle with spin half can exist in two states, one with spin up and the other with spin down. These two states, namely spin up and spin down, are designated by spins +1/2 and −1/2 respectively. The quantum physics necessary to understand the degenerate electron gas pressure was given by Wolfgang Pauli in 1925. It is called Pauli's exclusion principle. According to this principle, no two electrons in a given energy level can have the same spin configuration, i.e., both of them can not simultaneously have spin up or spin down. This tells us that a given energy level can be occupied by only two electrons, one having spin up and the other having spin down.

We have seen earlier (see Box 2.2) that the speeds of the particles in a classical ideal gas have a particular distribution called the Maxwell–Boltzmann distribution. The average speed of the particles decreases with decrease in

(continued)

Box 3.1 (continued)

temperature of the gas. As the temperature approaches absolute zero, the speeds of particles also approach zero value. As the speeds of the particles determine the pressure of the gas, the pressure of a classical ideal gas decreases with decrease in temperature. This can not happen in a quantum gas, as we will see below, because there are limits on the number of particles that can be present in any energy level.



A sketch of electrons occupying energy levels in a quantum gas is shown in the figure for three different values of temperature and density. The energy of the levels increases upwards, i.e., E_1, the energy of the bottommost level is lowest and E_6, the energy of the topmost level shown in the figure is highest. Note that the higher the energy of an electron, the higher is its speed. The panel on the left shows electron gas with particle density n and temperature T_1. The gas is not yet degenerate and particles are distributed over several energy levels. As the temperature decreases, the electrons occupy lower and lower levels as shown in the middle panel which shows electron distribution for gas having the same density n and temperature T_2, T_2 lower than T_1. When the density increases, the electrons start filling up lower levels, and even at low temperature some of them are forced to occupy higher levels. This is shown in the panel at the right which shows electron distribution for electron gas having an even lower temperature T_3 and density n'. Note that T_3, T_2, and T_1 are in increasing order and n' is higher than n. The filling up of higher energy levels results in high speeds of electrons for the gas shown in the right panel, even though its temperature (T_3) is lowest. Thus, its pressure remains high. The distribution of electrons in the energy levels can not change with further decrease in temperature, as no levels with lower energies are available for the electrons to occupy. The gas is degenerate and the pressure is independent of the temperature.

This stage in the life of a star comes after significant contraction of its central portion after its final cycle of nuclear fusion gets over. Because of the large energy released during contraction (due to the conversion of gravitational potential energy) the outer layers of the star are blown away and appear as a

Fig. 3.1 The Helix planetary nebula. It is at a distance of 700 light years from the Earth. The central star is in the process of becoming a white dwarf. **Credit:** ESO

cloud of gas surrounding the remnant of the star. Early astronomers thought such objects to be planets and these objects were called planetary nebulae. A picture of a planetary nebula is shown in Fig. 3.1. The remnant of the star after its outer layers have been thrown out is in hydrostatic equilibrium, its gravity supported by the pressure of the degenerate electron gas.

Such remnant stars were discovered much before the above theory was known. They were stars which (i) were white in appearance indicating their high temperature and (ii) had low luminosities in spite of their high temperatures, indicating that they were small in size (see Box 2.3). Their sizes were found to be much smaller, by about a factor of 100, as compared to those of typical stars like the Sun. They are believed to have been named White Dwarfs in 1922 by Willem J. Luyten, because of their white colour and small size. To give an example, after 4.5 billion years, when the Sun will become a white dwarf star, its radius will decrease 100 times and will be similar to the radius of the Earth. The density of the white dwarf material will be about 10^9 kg m^{-3} i.e., a million times the average density of the Sun at present. The typical density of white dwarfs of different masses can vary between 10^7 to 10^9 kg m^{-3}. The temperature of a white dwarf keeps decreasing as it radiates away its thermal energy, and ultimately it might reach close to absolute zero. Then the star will not emit any radiation and will not be visible, and it would be more appropriate to call it a black dwarf. However, the cooling rate of white dwarfs, though high initially, slows down and it has been estimated that even after 10^{10} years of their formation, their temperatures would still be about 6500 K. White

dwarfs are thus stars whose gravity is balanced by the pressure of degenerate electron gas and are the end stages for stars having initial masses less than 8 $M_\odot$.

3.2.2 Award Winning Work: Upper Limit on the Mass of White Dwarf Stars

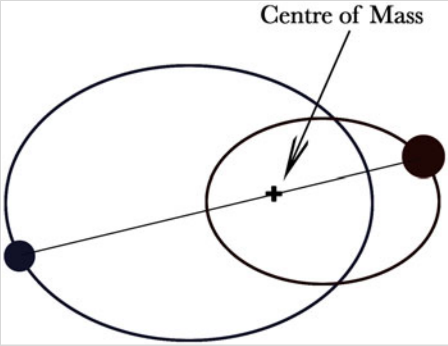
Two of the early stars to be identified as white dwarfs were Eridani B and Sirius B. They were first observed in 1783 and 1862, respectively. They were both members of binary systems. A binary system is a system of two stars gravitationally bound to, and going round each other (see Box 3.2). Sirius B is the companion of Sirius A which is the brightest star in the night sky. An application of Newton's laws of motion and his law of gravity to the components of a binary system can lead to the determination of their masses. The temperatures of the stars can be determined from their spectra assuming the stars to be emitting like black bodies, and the sizes of the stars can be determined from their temperatures and luminosities (see Box 2.3). All these measurements were made for Sirius B in the first two decades of the twentieth century and it was immediately realized that the star had enormous density, about 10^9 kg m^{-3}.

Box 3.2 Binary Stars and Determination of Their Masses

Almost half of the stars that we see in the night sky are members of binary systems. A binary system is a system of two stars, called binary stars, which are gravitationally bound to each other. They move in elliptical orbits around their centre of mass which is the point through which the total mass of the system effectively acts. The centre of mass is at the common focus of the two ellipses. A typical binary system is shown in the figure. If the two spherical stars in a binary system have equal masses, then the centre of mass will be at a point exactly midway on the line joining the centres of the two stars and the two stars will revolve in a circular orbit centred on that point. If one of the stars is heavier than the other, the centre of mass will shift towards the heavier star along that line and the orbits of the stars will in general be elliptical, with the centre of mass at one of the foci of the ellipses. Note that, during their motion, the two stars are always on the opposite sides of the centre of mass.

(continued)

Box 3.2 (continued)



Just like the planetary motions follow Kepler's laws, the motion of members of the binary also obey Kepler's laws. The time period of orbital motion of the stars (which is same for both stars) depends on the sum of masses of the two stars and the average distance between them. If we can measure the period and the average distance between them, we can obtain the total mass of the stars. The ratio of the distances of the two stars from their centre of mass directly gives the ratio of the two masses. By studying the orbits of the stars individually we can locate the position of the centre of mass and thus, obtain the distances of stars from it. From these distances, we can get the ratio of the masses of the two stars. This ratio and the total mass then yield the masses of the individual stars. Even though the determination of masses seems possible for any binary system, in practice it is not always so. The main reason for this is that the orbits of the binary stars may not be in a plane perpendicular to the line of sight to these stars. For this reason, the parameters of the orbits that we can measure are only their projected values and will not yield correct values of the masses.

Even though the details of stellar evolution as described in the previous section were not known in the early twentieth century (that nuclear reactions are the source of energy inside the stars was firmly established by Bethe only in late 1930s), it was clear that white dwarf stars have very high densities and must be collapsed stars. It was realized that at such high densities, the electrons are free and form an important component of the gas, which was assumed to be a classical ideal gas at that time. The large densities of the stars posed a theoretical problem and gave rise to a paradox which was pointed out by Eddington. He wondered what would happen to such a star when it radiated away its thermal energy and cooled. The nuclei would then have to recombine with electrons, forming atoms as that would be the favoured state at lower temperatures. Atoms have a much larger volume as compared to bare nuclei and can not be accommodated in the small volume of the white dwarf. The star

would then have to expand. However, its lower pressure (at lower temperature) would not allow for such expansion to take place. Eddington summarized the situation as "the star is continuously losing energy but has insufficient energy to grow cold."

The British astronomer R. H. Fowler realized that at the high densities in white dwarf stars, the electron gas would no longer remain a classical ideal gas, rather quantum mechanical effects would become important and the electron gas would become degenerate. He calculated the pressure of the degenerate electron gas and used it, instead of the pressure of classical ideal gas, to calculate the structure of white dwarf stars. He showed in 1926 that the paradox posed by Eddington no longer existed. According to his calculations the radius of a white dwarf decreases with increasing mass, and white dwarfs of all masses can exist in hydrostatic equilibrium. It is worth noting that the quantum mechanical results necessary to perform the calculation of degenerate gas pressure were published only that year.

Chandrasekhar in 1930 realized that Fowler's calculations missed the fact that at the high densities in the white dwarfs, the electrons would be forced to occupy energy levels with energies high enough (see Box 3.1) so that their speeds will be close to the speed of light. For electrons having such relativistic velocities, the calculation of gas pressure has to be modified as Newton's laws of motion no longer apply and the results of special theory of relativity have to be used (see Box 2.5), i.e., the relativistic effects have to be taken into account. Chandrasekhar took into account these effects and calculated the degenerate electron gas pressure in white dwarfs. He then used it to calculate the structure of white dwarf stars. He did this work on a ship on the way to Cambridge where he was going just after his graduation, to pursue his Ph.D. His calculations led to a very surprising result. It turned out that there is an upper limit to the mass of a star which can be supported by the pressure of degenerate electron gas, i.e., there is an upper limit to the mass of a white dwarf star. He further improved on his results after arriving in Cambridge and obtained a limit on the mass of a white dwarf. This limit is known as the Chandrasekhar limit. Its exact value was later calculated to be 1.44 $M_{\odot}$. Remember that this is the end stage of all stars which start with an initial mass smaller than 8 $M_{\odot}$. The stars lose their outer material which forms a planetary nebula, and the Chandrasekhar limit applies to the mass of the stellar remnant which is the white dwarf.

Chandrasekhar presented his results in the meeting of the Royal Society of England in January, 1935. Eddington, who was very famous and influential, ridiculed the idea of such a limit as it indicated that stars with masses above this limit would not be able to achieve equilibrium and will be doomed to

contract for ever. Eddington found this absurd and commented that nature could not behave in such an absurd fashion. This delayed the acceptance of Chandrasekhar's work. He was awarded a Nobel Prize only in 1983. Note that all known white dwarfs to date have masses smaller than the Chandrasekhar limit.

3.3 Nobel Prize 1983: W. A. Fowler

The Nobel Prize for the year 1983 was given to William A. Fowler "for his theoretical and experimental studies of the nuclear reactions of importance in the formation of the chemical elements in the Universe." He shared the prize with S. Chandrasekhar.

William Fowler was born in 1911 in the USA. He graduated from The Ohio State University and later obtained his Ph.D. at the age of 25 from the California Institute of Technology, Pasadena (Caltech). He continued to work there as a faculty, becoming a professor in 1946. During the world war he conducted military research. Later he became the director of the Kellogg Radiation Laboratory there.
Fowler spent sabbatical years in Cambridge, England and collaborated extensively with the famous astronomers Fred Hoyle, Margaret Burbidge, and Geoffrey Burbidge on the synthesis of elements in stars. With Hoyle, he also worked on problems in radio astronomy. He became very interested in English culture and perhaps was planning to settle in England, which did not materialize. He retired in 1982 and passed away in 1995.

3.3.1 Background

We have seen that stars having their initial mass smaller than about 8 $M_\odot$ end up as white dwarfs. What happens to the stars above this mass limit? These stars undergo further nuclear fusion cycles until the core material is converted into iron. No nuclear energy can be generated beyond this stage through nuclear fusion as the mass of the reactants happens to be smaller than the mass of the products and energy has to be supplied if such reactions are to take place. In absence of any energy generating nuclear reactions, the process of gravity

taking over and contraction of the star follows. The contraction continues even after the electron gas becomes degenerate, as the gravity is simply too large for the pressure of degenerate electron gas to balance and the star can not find equilibrium as a white dwarf. Thus, the density and temperature at the centre keep increasing until the collapsing object reaches a radius of a few tens of km and the density of the stellar material becomes extremely high–about 10^{17} kg m^{-3}.

The nuclei of iron and those of other elements which may be present in the dense core of the star, cannot sustain themselves under these densities and temperatures in the core of the star; they disintegrate into their constituent particles, i.e., the neutrons and protons. Due to the high density, the protons and electrons are very close to one another and combine together to form free neutrons emitting an electron neutrino in the process as per the following reaction.

$$\mathrm{p} + \mathrm{e}^- \rightarrow \mathrm{n} + \nu_\mathrm{e}$$

Here n represents a neutron. It may be noted that the mass of a neutron is larger than the mass of a proton. The above reaction is only possible as the electrons in the degenerate gas have extremely high energies which can generate the necessary mass to create a neutron from a proton. Extremely large amounts of energy, which can be greater than several times the energy given out by an entire galaxy in a second, is generated during the collapse of the core of the star at the expense of its gravitational potential energy. A large number of neutrinos are generated during the conversion of protons and electrons to neutrons. These also exert pressure on the outer stellar material. The sudden release of a large amount of energy and a large number of neutrinos results in a gigantic explosion in which the outer parts of the star are thrown out in a spectacular fashion. Such a stellar explosion is called a core-collapse supernova or Type II supernova.

We can get an idea of the dynamics involved in such explosions from the following. In 1054 AD, it was noted in Chinese chronicles that a new star which was called a "guest star" suddenly appeared in the sky. For several days the star remained bright enough to be seen during day time too, before fading out. Mention of the guest star has also been found in Japanese records. Records of its sighting are also found in a thirteenth century Arabic text. While the Chinese could see the object that time with their naked eyes, we now need a telescope to observe the object at the same position in the sky. This object, shown in Fig. 3.2, has been named the Crab nebula because of its appearance. What we see there are gases which are still dispersing as a result of the stellar explosion that was seen by the Chinese astronomers. These gaseous filaments

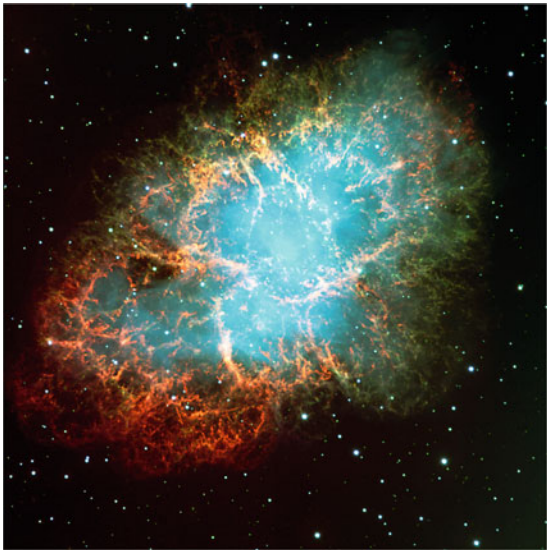


Fig. 3.2 The Crab nebula, which is the remnant of the supernova explosion seen by the Chinese in 1054 AD. What we see are the gases still expanding with velocities exceeding 1000 km s^{-1} as a result of the explosion. **Credit:** ESO

are moving with velocities of about a few thousand km s^{-1} after more than a thousand years of the stellar explosion. We can recall that the fastest speed we, at times, drive on highways is not too much higher than 100 km per hour! Thus, one can get some idea of the velocities and accelerations that these gases must have had at the time of the explosion. The structure that we see now is known as a supernova remnant, which constitutes the remains of a supernova explosion. While the supernova was observed on the Earth in 1054 AD, it actually happened about 6523 years before that date, which we know since the distance to the site of the explosion is known to be about 6523 light years from us. Theoretical studies suggest that the progenitor star which exploded as this supernova, had initial mass (before explosion) in the range 8–11 solar masses.

Box 3.3 Supernova 1987A

On February 24, 1987, Canadian astronomer Ian Shelton of the University of Toronto took a picture of the Large Magellanic Cloud at the Las Campanas observatory in Chile through a telescope. To his surprise, the photograph showed an extremely bright star that was not present in the photographs taken earlier. Just to confirm that it was a real star and not merely an artifact or a defect in the photographic plate, he went outdoors and looked at the Magellanic cloud with his naked eyes. He could see the new star clearly. He immediately realized that the

(continued)

Box 3.3 (continued)
new star was actually a supernova explosion. The supernova was independently discovered on the same day by Oscar Duhalde in the same observatory.
With the help of observations taken by the International Ultraviolet Explorer satellite astronomers identified that a blue supergiant having a mass of about 20 times the mass of the Sun and a radius of about 40 times the radius of the Sun (hence giant, the blue colour indicating high temperature) was earlier present at the location of the exploding star. They named the exploding star Supernova 1987A. The letter A indicating that it was the first supernova seen in the year 1987. Astronomers believe the original star swelled up (which lowered its temperature) to become a red supergiant, shed some of its mass, and then contracted and reheated to become a blue supergiant. In less than a second, the star's core collapsed, and the released neutrinos heated the inner core to 10 billion degrees. The process tore the star apart, releasing a burst of neutrinos into space. Shelton was extremely lucky to have been observing at the right time to actually observe the exploding star. A few hours before Shelton saw it, several neutrino observatories including the Super-Kamiokande in Japan (see Sect. 2.3.2), had recorded a sudden large flux of neutrinos. It has been concluded since that these neutrinos came from supernova 1987A.

3.3.2 Award Winning Work: Formation of Chemical Elements in the Universe

How, where, and when are the chemical elements that we see around us, produced and that too in the relative amounts in which they are observed? Bethe in 1939 had considered the possibility that elements heavier than helium could be generated inside stars by adding protons one by one to the helium nuclei which were produced in the hydrogen burning phase. He showed that this process can not work in stellar interiors, as all elements up to boron disintegrate on combining with protons, and suggested that the heavy elements found in stars must have already been there when a star formed. The process of formation of elements in stars is called stellar nucleosynthesis. The other process which could synthesize elements is called the big bang nucleosynthesis. It is believed to have occurred in the first few minutes in the life of the Universe according to the big bang theory of the Universe about which we will learn in details in Chap. 8. Here, it will suffice to know that at the time of big bang nucleosynthesis, the Universe was completely homogeneous and uniform. It was almost entirely made up of elementary particles and extremely high energy radiation, and its temperature and density were high enough for nuclear reactions to take place. The density and temperature slowly decreased with time and after about 20 minutes in the life of the Universe, the

nuclear reactions ceased. Big bang nucleosynthesis took place during these few minutes.

In order to calculate the quantities of different elements produced during the stellar and big bang nucleosynthesis, one needs accurate knowledge of the rates at which different nuclear reactions take place. This information comes through experiments as well as from theoretical calculations. Knowing these rates, one has to perform detailed calculations of stellar structure and evolution and of the evolution of the Universe. William Fowler was an experimental nuclear physicist and together with his numerous collaborators, he also carried out theoretical calculations including those of stellar and big bang nucleosynthesis. He made important contributions to all aspects of nucleosynthesis. Below, we will look at some of his important contributions.

Early estimates of the big bang nucleosynthesis were presented by George Gamow and his collaborators in 1948, suggesting that all known elements could be synthesized in roughly the required (observed) proportions in this process. In the big bang nucleosynthesis the chemical elements would have to be synthesized one by one through the capture of one neutron at a time, which would then get converted to a proton by emitting an electron and an antineutrino (antiparticle of an electron neutrino), thereby producing the nucleus of a heavier element, i.e., an element with higher mass number. Such reactions can take place at a higher rate compared to reactions between two protons or between two nuclei due to the absence of electrostatic repulsion, the neutron being chargeless. A nucleus thus produced would in turn capture another neutron and the process would go on to produce nuclei of heavier and heavier elements. For a heavier nucleus to be produced, the reacting nucleus has to be stable at least until a neutron can be captured. However, problems existed for big bang nucleosynthesis as well, as no nuclei with mass numbers 5 and 8 exist in nature. This is known as the mass gap at mass numbers 5 and 8 and indicates instability of such elements. Such nuclei, i.e., ^{5}He, ^{5}Li and ^{8}Be could of course be produced in the big bang nucleosynthesis; however, they were found to be highly unstable and decayed almost instantly into nuclei of lighter elements. The instability of ^{8}Be was experimentally verified by Fowler in 1949. Fowler's results, along with experiments by others, proved that no stable nuclei at the mass gaps exist, which prohibits the build up of heavier elements in the big bang nucleosynthesis. Fowler in 1967 performed big bang nucleosynthesis calculations using the then latest data on the rates of nuclear reactions, and was able to produce the right amounts (i.e., consistent with their observed abundances) of ^{2}D, ^{3}He, ^{4}He. Some amount of ^{7}Li could also be produced. The first generation of stars thus, had only these light elements.

The idea that all heavy elements present in the Universe were synthesized in stars was put forth by Fred Hoyle in papers published in 1946 and 1954. The important reaction in this process was the triple alpha reaction to produce carbon (see Sect. 2.2.2). For Hoyle's ideas to work, this reaction would have to take place at temperatures just above 10^8 K, while theoretical calculations using the then available reaction rates showed that the reaction would require temperatures of 2×10^8 K. This was because the probability of three alpha particles to come close enough to combine was very small, considering the positive charge of two units of each alpha particle. Hoyle suggested that the carbon nucleus must have an excited state with the energy of 6.65 MeV (MeV which stands for million electron volts, is a unit of energy). The presence of such an excited state would greatly enhance the theoretically calculated rate of the triple alpha reaction, which can then take place at lower temperature and allow for Hoyle's scheme of stellar nucleosynthesis to work. Hoyle visited Fowler's laboratory and discussed his idea and suggested that experimentalists should be able to confirm the presence of such a state. Fowler and his collaborators performed experiments to study the decay of ^{12}Be, and showed in 1957 that it leads to the excited state of ^{12}C (as predicted by Hoyle), which then decays to three alpha particles. Nuclear reactions have a property of reversibility according to which if a particular nuclear reaction can take place then its reverse reaction, i.e., the products reacting and producing the reactants, can also take place. According to this, Fowler's experiment showed that the reverse reaction, i.e., the triple alpha reaction is indeed possible at the required rate and can produce carbon in its excited state. Even though the ^{12}C thus produced in the excited state most often breaks back into three alpha particles, there is a finite probability, of about one in 2500, that it goes to the stable ground state of ^{12}C, thereby overcoming the mass gaps. This proved that Hoyle's speculation was indeed correct and provided crucial support to Hoyle's ideas that the synthesis of elements heavier than helium can indeed take place inside stars. It is believed that most of the observed lithium, and all of beryllium and boron seen in the Universe are produced through the interaction of cosmic rays (see Box 2.9) with interstellar matter, namely the matter in the space between stars.

We have seen (Sects. 3.2.1 and 3.3.1) that elements up to iron are produced inside stars. Fusion reactions involving iron do not generate energy and in fact require energy to be supplied for them to take place. How then are the elements heavier than iron synthesized? These are mostly produced by what are known as the s-process and the r-process, in which neutrons are captured by nuclei one by one. The neutrons decay into protons producing nuclei of heavier element. The letters s and r stand for slow and rapid neutron capture

rates, as compared to the decay rates of the generated nuclei. The s-process takes place when the neutron densities are small, about 10^6 neutrons cm^{-3}. Such densities are present in the envelopes of stars with masses in the range of 1–10 $M_\odot$ when they are in late stages of evolution. These reactions produce about the first half of the elements heavier than iron. The r-process takes place when the density of neutrons is large, about 10^{23} neutrons cm^{-3}. Such densities are only available in the exploding material of supernovae, where a large number of neutrons are produced as seen in Sect. 3.3.1. These r-process reactions produce the remaining heavy nuclei all the way up to uranium. Fowler made significant contributions to both these aspects of nucleosynthesis. Recently, heavy elements were seen in material surrounding a merged neutron star binary which was detected through gravitational waves which will be discussed in Chap. 7.

The heavy elements produced in stars are thrown out in the stellar explosions like the one's responsible for formation of planetary nebulae and also in supernova explosions. These then enrich the interstellar matter. New stars formed from this matter have a higher amount of heavy elements. The abundances of heavy elements in the Universe thus increase with time through these processes. This has been confirmed observationally.

In 1957, Fowler with his collaborators Geoffrey Burbidge, Margaret Burbidge, and Fred Hoyle wrote a seminal paper entitled "Synthesis of elements in stars" in which they presented results of stellar nucleosynthesis starting from helium all the way up to uranium. This paper is referred to as B^2FH and is considered to be a landmark paper in stellar nucleosynthesis. This paper reviewed the entire nucleosynthesis scenario. Fowler's theoretical and experimental results were some of the key inputs to this paper.

Fowler also worked on the rate of neutrino reactions with chlorine. He considered all neutrinos produced in the PP chain and showed that a small fraction of these neutrinos should be energetic enough to be detectable through interaction with chlorine nuclei. He conveyed this to Davis and prompted him to set up his experiment (see Sect. 2.3.2). Fowler also experimentally determined the reaction rates of most of the reactions in the CNO cycle proposed by Bethe, which was important for accurate calculations of stellar structure and evolution.

Fowler was awarded the Nobel Prize for his enormous contribution towards the understanding of the synthesis of elements in the Universe.

4

Radio and X-Ray Astronomy

4.1 Introduction

The Nobel Prizes in the category of radio and X-ray astronomy, were awarded in 1974 and 2002 to three astronomers: Martin Ryle, Anthony Hewish, and Riccardo Giacconi. Their work involves observations of cosmic objects in different bands of the electromagnetic spectrum. The award for 1974 was shared by Ryle and Hewish and was in the field of radio astronomy. This was the first Nobel Prize awarded to professional astronomers. The award to Ryle was for developing the field of radio astronomy and the techniques used therein, while that to Hewish was for the detection of the first pulsar. Ryle's work started in the 1940s while Hewish started his work in 1965 and the discovery of pulsar was made in 1967. The Award to Giacconi was in the field of X-ray astronomy, for his significant contribution to the field culminating in the discovery of X-ray sources. His work on X-ray astronomy started in the 1950s and continued for several decades.

4.2 Nobel Prize 1974: M. Ryle

The Nobel Prize for the year 1974 was given to Martin Ryle "for pioneering research in radio astrophysics: observations and inventions, in particular of the aperture synthesis technique." He shared the prize with Anthony Hewish.

P. Khare, *Nobel Prizes in Astronomy*, Astronomers' Universe,
https://doi.org/10.1007/978-3-031-29639-0_4

Credit: AIP Emilio Segrè Visual Archives, John Irwin Slide Collection

Martin Ryle was born in 1918 in England. He studied physics at Christ Church, Oxford from 1936 to 1939. During the war years he worked in the Telecommunications Research Establishment, on the design of antennas for radar equipment aboard an aircraft. After the war, he received a fellowship at Cavendish Laboratory of Cambridge University. In 1948 he was appointed a lecturer in physics at Cambridge and later, in 1959 he became the Chair of radio astronomy. In 1957, he was appointed the founder director of the Mullard Radio Astronomy Observatory. Under his leadership, the Cambridge group produced the Cambridge Catalogue of Radio Sources. The third Cambridge catalogue (3C) remains very useful even today.

Ryle was an outspoken critic of nuclear power and nuclear weapons. From 1976 onward he worked passionately towards promoting socially responsible use of science and technology. One of his famous quotes at the end of a letter that he wrote and which was published posthumously read "Our cleverness has grown prodigiously—but not our wisdom." He passed away in 1984.

4.2.1 Background

Astronomy is considered to be one of the oldest sciences. Till the seventeenth century all observations of heavenly objects were made with the naked eye. Detailed observations of positions of stars and planets were made by several observers in this period. On the basis of such observations, Copernicus proposed the heliocentric model of solar system in his book published in 1543. About half a century later, Kepler gave his three laws of planetary motion in his book published in 1609. In the same year, Galileo started observing heavenly objects through a telescope that he had made himself. The year 1609 is therefore considered to be the birth of modern astronomy.

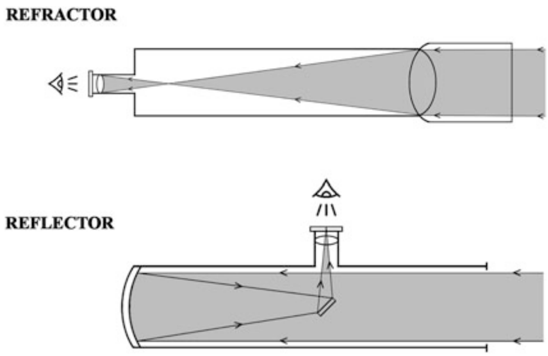


Fig. 4.1 Sketches of refractor (upper panel) and reflector (lower panel) optical telescopes. These use lenses and mirrors, respectively

The main purpose of a telescope is to collect light. Galileo's telescope used lenses. Such telescopes collect light from a distant object that falls on the front lens to a focal point. As the front lenses used in telescopes have bigger diameter than our pupils, more light falls on them as compared to our eyes and is collected at a point. As a result, when we look at an object through a telescope, we see the object as being brighter. Thus, with the aid of a telescope, we can see fainter, and therefore farther, objects. Clearly, the bigger the diameter of the lens (the cross-sections of lenses used in telescopes are circular), the larger will be the amount of light that falls on it and is collected at the focus. Newton for the first time used a concave mirror instead of a lens, which also focuses the light falling on it at its focal point. Telescopes using lenses are called refracting type and those using mirrors are called reflecting type. Sketches of the two types are shown in Fig. 4.1.

Big mirrors are much easier to make and handle than big lenses and also have less optical defects in their images. Thus, all modern telescopes made in the twentieth century have used mirrors. The largest mirrors used in telescopes to date have a diameter of 10 m. Telescopes with bigger mirrors are under construction, the notable one being the Extremely Large Telescope being constructed by the European Space Observatory with a mirror of 39 m diameter. All the telescopes discussed above are used to observe heavenly objects in visible light (see Box 2.1), that they are known to emit and which our eyes can see. These are called optical telescopes. The mirrors used in optical telescopes have to be extremely smooth, any permissible roughness should have a size smaller than the wavelength of visible light– which is smaller than a millionth of a metre– in order to produce an accurate image of the object. The mirrors thus have to be highly polished.

In 1930s, engineers in the Bell Labs, USA were experimenting with voice communication with the help of radio waves using an antenna. Their signals

Fig. 4.2 Full size replica of the radio antenna used by Jansky, with which radio waves coming from astronomical sources were detected for the first time. This was the world's first radio telescope. **Credit:** NRAO/AU/NSF

were getting disturbed by a noisy static which was interfering with the shortwaves (which are radio waves having a wavelength smaller than 200 m) that they were using for the purpose. Karl Jansky was given the task of investigating the causes of the static. After detailed study, he discovered that the direction of the source of noise slowly shifted across the sky with time during the day. It was clear to him that the noise was not coming from the Sun but from somewhere else. He was an engineer and had no knowledge of astronomical objects. He then consulted an astronomer, and after detailed study came to the conclusion that the source was outside the solar system and the maximum noise was coming from the direction of the centre of the Milky Way. He published a paper entitled "Radio Waves from Outside the solar System" in 1933. This is considered to be one of the most important papers in astronomy. It established that cosmic sources can emit radio waves in addition to visible light. This single observation can be considered to be responsible for the birth of radio astronomy. The antenna used by Jansky can thus be considered to be the first radio telescope. A full size replica of that antenna, located in the National Radio Astronomical Observatory in West Virginia, USA, is shown in Fig. 4.2.

A radio telescope is a specialized radio antenna combined with a radio receiver. Radio waves have a very wide range of wavelengths starting from a cm to several km. As a result, radio antennas (telescopes) have widely different designs, shapes and sizes. Telescopes used to observe/receive waves having long wavelengths, i.e., between 3 to 30 m, have directional antennas like the TV antennas used inside homes. For waves having shorter wavelengths, the antenna is in the form of one or more parabolic dishes, which are essentially

Fig. 4.3 Picture showing a few of the 30 dishes of the Giant Metrewave Radio Telescope (GMRT) near Pune, India

mirrors for radio waves and focus the radio waves coming from an object at its focal point where a radio receiver is mounted.

Observing with radio telescopes has some advantages over optical telescopes. Radio telescopes can be used to observe astronomical objects during the day as well, unlike optical telescopes (except solar telescopes which observe the Sun), which can only be used at night. This is because visible light from the Sun gets scattered in all directions by the dust and molecules in the Earth's atmosphere and as a result, the whole sky appears bright and we can not see stars and galaxies during the day. Radio waves do not scatter significantly due to their longer wavelengths and as a result, the radio sky, i.e., the sky when viewed through radio waves, is dark even during the day, and we can observe the radio waves coming from any cosmic radio source as long as it is not lying directly behind the Sun. Also, unlike visible radiation, radio waves are not absorbed (and scattered) significantly by interstellar matter, and hence radio telescopes can observe much farther sources.

Shortwave radio telescopes differ from optical telescopes mainly in two respects. The first is that, because of the much longer wavelengths of the radio waves compared to those of visible radiation, the radio mirrors, i.e., dishes do not have to be as smooth as the mirrors used in optical telescopes. In fact these dishes are often made of wire meshes, which reduces their cost and their weight and makes them easier to handle. A picture of a few dishes of the Giant Metrewave Radio Telescope located near Pune, India is shown in Fig. 4.3.

In order to understand the second and major difference between the two types of telescopes, we have to understand another important property of a

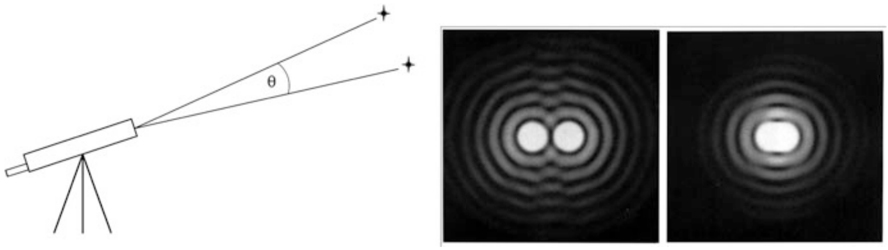


Fig. 4.4 The leftmost panel shows a telescope observing two stars having angular separation of θ. The two other panels show the images produced by the telescope for angular separation slightly lower than (rightmost panel), and slightly higher than (middle panel) the inverse of the resolving power of the telescope

telescope: its resolving power. It is the measure of the ability of a telescope to produce separate images of two nearby objects or sources. Thus, it is the inverse of the smallest angle α between two objects whose images produced by the telescope are distinct. This is shown in Fig. 4.4, the middle and right panels of which show images of two stars close to each other for two different values of angular separation. Angular separation is the angle θ, between the directions of the two stars as seen from the site of the telescope as shown in the left panel of the figure. In the rightmost image, the angular separation between the two stars is slightly smaller than α, the inverse of the resolving power of the telescope. Thus, the images of the two stars are not distinct. In the image in the middle, the angular separation between the stars is slightly larger than α which is why the images produced by the telescope are distinct. The higher the resolving power of a telescope, the better can it resolve two nearby objects in the sky.

The resolving power of a telescope depends on the ratio of the diameter of the telescope mirror/dish to the wavelength of the radiation observed. The larger the mirror, the smaller is the angular separation between two sources that the telescope is able to resolve, i.e., the higher is its resolving power. As we know, radio wavelengths are at least ten thousand times larger than optical wavelengths. Thus, in order for a radio telescope to have the same resolving power as an optical telescope, the diameter of its dish should be more than ten thousand times as large as that of the mirror of an optical telescope. A typical ground based optical telescope has resolution of about 1 arc sec (arc second, which is equal to 1/3600 of a degree). To have resolving power similar to even this telescope, the diameter of the dish of a radio telescope observing at 10 cm will have to be about 20 km. Obviously this is an impossible feat to achieve.

4.2.2 Award Winning Work: Aperture Synthesis Technique in Radio Astronomy

In Cambridge, Ryle first started to study the radio properties of the Sun. At that time, the available antennas could not even resolve a source as big as the Sun. Ryle soon decided to address the problem and directed his group to work towards improving the methods then used to observe radio sources. This led to two new techniques which are now used widely in radio astronomy namely, interferometry and aperture synthesis, which are described below. We have seen above that radio telescopes having dishes of a few meter diameter have much lower resolution than even small optical telescopes. One way to get higher resolution for a radio telescope is to use interference (see Box 4.1). Australian astronomers were the first to use this technique. Using a single antenna, they observed light from the Sun directly as well as that reflected by sea. Ryle used two dishes and combined the signal from the two, taking care to allow for the time difference between the arrival of sunlight at the two dishes. The resolving power in this case is determined by the distance between the two dishes rather than the diameter of these dishes.

Box 4.1 Interference

Interference is a wave phenomenon. When two waves having the same wavelength and traveling in the same direction superimpose, they can produce interesting effects like destroying each other (destructive interference) or augmenting each other (constructive interference), or can have an in-between effect depending on the way they superimpose. An example of constructive and destructive interference is shown in Fig. A, where the resultant wave obtained by the superposition of wave 1 and wave 2 is shown at the bottom for two different cases.

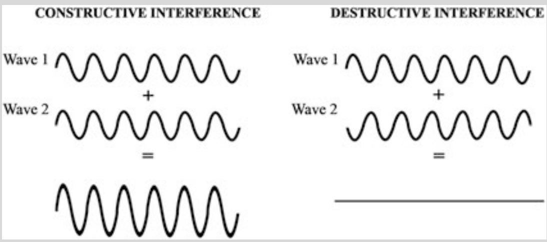


Fig. A

(continued)

Box 4.1 (continued)

For light, the effect was experimentally studied for the first time by Thomas Young through his famous double slit experiment. He passed light from a single source through two narrow slits separated by a small distance, in an opaque wall, effectively creating two sources of light having identical properties. The light from both slits was made to fall on a screen as shown in Fig. B. The two waves have to travel different distances to reach any given point on the screen (except for the middle point O). Thus, the brightness produced at that point depends on the difference in the path lengths traveled by the two beams. For example at the central point the path difference is zero and the waves interact constructively, as was seen in the left panel in Fig. A, producing a bright band (image of the slits) there. Similarly, bright bands will be present at points on the screen where the path differences between the two beams are integral multiples of the wavelength so that the two waves will interfere constructively as in the left panel of Fig. A. At points where the differences in path lengths are odd multiples of half wavelength, the waves will combine destructively as was seen in the right panel of Fig. A, producing dark bands there. The interference thus results in alternate bright and dark bands, called fringes, on the screen as shown in the rightmost strip in Fig. B. Note that the screen has been turned through 90° in this strip to make the fringes visible to the reader. In absence of the wall having slits, the screen would have been uniformly illuminated. The effect seen on the screen is because of the interference of the waves coming from the two slits. The pattern of fringes, i.e., the distribution of intensity produced on the screen is known as interference pattern. An interferometer is essentially an instrument which generates interference in waves coming from the same source but along different paths like in Young's double slit experiment as described here. We will look at interferometers in greater detail in Chap. 7.

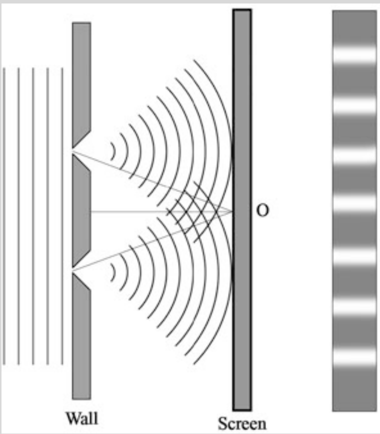


Fig. B

In 1946, Ryle along with his collaborator Derek Vonberg, constructed a radio interferometer using two radio antennas separated by a few metres. The signals from a source, obtained in the two antennas were combined. By studying the resulting interference pattern, they could improve the resolution. They used the technique to observe the Sun and were able to show that the radio emission coming from the Sun originated from a region which had the same area as that of a big Sunspot that had appeared on the Sun in the same year. They were the first to publish a paper based on observations made using more than one antenna. This could perhaps be hailed as the birth of radio interferometry.

Using interferometry, Ryle then developed the technique of aperture synthesis. This is a type of interferometry that combines signals from a number of radio telescopes/dishes to produce an image which appears to have been produced by a single telescope having much larger aperture (dish diameter) and having much higher resolution than those of images produced by individual telescopes/dishes. Ryle initially used two dishes. He mounted them on rails which allowed them to be moved up to a maximum distance of 1.6 km between them. The straight line distance between the centres of the two dishes as seen from the source, is called baseline. He took observations of a source for several values of separations between the two antennas, i.e., for several baselines of different lengths. Combining and analyzing the data, he obtained a resolution which was comparable to that of a single telescope having a dish diameter of 1.6 km. Thus, he had synthesized a large effective aperture by combining two smaller aperture telescopes. For each baseline, the data from the antennas produce an output called complex visibility, which is a single mathematical component of the actual image, called the Fourier component. Several Fourier components of an image, obtained by using different baselines, can be combined to reconstruct the image. The larger the number of Fourier components, the more accurate is the constructed image. Radio sources are often extended objects and the image of a source is also called the radio map of the observed object. If one uses more than two antennas, each pair of antennas, for each value of the separation between the two, produces one Fourier component of the image. For a group of N antennas the number of pairs is $N(N-1)/2$, yielding that many components of the image (even more if the distances between the antennas could be varied). All these components can be combined to get a very accurate map of a radio source.

Moving the bulky radio antennas can be quite laborious and is not sustainable in the long run. So instead, Ryle took advantage of the Earth's rotation around its axis. The sources rise from the east and set in the west. The antennas are rotated to track a source. Because of this the length of the baseline changes,

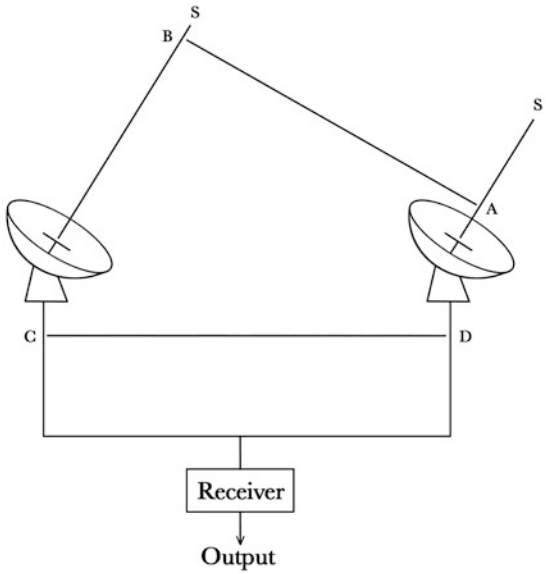


Fig. 4.5 Change in the length of baseline due to the rotation of the Earth. At the position of the antenna shown in the figure, S indicates the direction of the source and the baseline is AB. When the source is at the zenith, the baseline is CD. Thus, the length of the baseline changes continuously as the antennas track the source during observation, giving us a large number of Fourier components of the image which yield a high resolution image

as explained in Fig. 4.5. Thus, observing a source over some period of time yields observations for several baselines of different lengths, and increases the number of Fourier components of the image that are obtained. This results in enhancing the resolution and accuracy of the map obtained. Complex electronics is used in order to combine the outputs of various antennas and complex computer analysis is needed to get the radio map from the data so obtained. Ryle was awarded the Nobel Prize for his enormous contribution towards developing the techniques used widely by Radio astronomers.

Box 4.2 Some Big Radio Telescopes

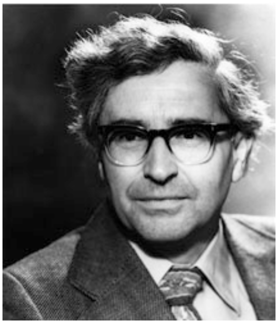
The largest single dish radio telescope is the Five-hundred-metre Aperture Spherical Telescope (FAST). It has a diameter of 500 m and is located in a natural depression in Guizhou district of China. It was completed in 2016.
Modern radio telescopes use several antennas collectively called an array. These are arranged in such a way that the Fourier components obtained are best suited to enhance the accuracy of the radio map. The Giant Metrewave Radio Telescope in India has 35 dishes, each having a diameter of 45 m. Fourteen dishes are spread randomly over a central region of one square km area and the rest are arranged in an approximate Y shape, each arm having a length of 3 km. Some of these dishes can be seen in Fig. 4.3. The Very Large Array (VLA) in New Mexico has 27

(continued)

Box 4.2 (continued)
antennas having a diameter of 25 m each. The antennas are mounted on rails and can be moved to generate more baselines.
Over time, owing to advances in technology, radio interferometry has made rapid strides. Interferometers have dishes with large distances between them. We have the Very Large Baseline Interferometer (VLBI) which has been operating since the late 1970s and combines the signals obtained by several dishes located in different parts of the world. This is thus equivalent to observation by a single dish radio telescope with diameter of thousands of km. One of the biggest arrays is the Event Horizon Telescope, which uses several existing telescopes and has been able to achieve an angular resolution of 43 micro arc sec, far superior to the highest resolution obtained by optical telescopes which is 0.002 arc sec for the Very Large Telescope (VLT) in Chile. The most ambitious radio telescope, which is under construction at present, is the Square Kilometre Array (SKA) which will have two arrays of antennas, one in South Africa and the other in Australia. This will have a light collecting area of 1 square km and will further enhance the resolution obtained by radio telescopes.

4.3 Nobel Prize 1974: A. Hewish

The Nobel Prize for the year 1974 was given to Anthony Hewish "for pioneering research in radio astrophysics: for his decisive role in the discovery of pulsars." He shared the prize with Martin Ryle.



Credit: AIP Emilio Segrè Visual Archives, Weber Collection, W. F. Meggers Gallery of Nobel Laureates Collection

Anthony Hewish was born in England in 1924. During his graduation itself, he had to join the war service during which he worked with Martin Ryle. After the war, he completed his graduation at Cambridge University, and obtained

his Ph.D. in 1952 at the Cavendish laboratory working with Ryle and his team. He served as a professor at the Cavendish laboratory from 1971 to 1989. During 1982 to 1988 he was also the head of the Mullard Radio Astronomy Observatory, Cambridge. He had been a Professor at the Royal Institution in London since 1977. He passed away in 2021.

4.3.1 Background

We have seen in Sect. 3.2.1 that stars with initial mass smaller than about 8 $M_\odot$ end their lives as white dwarfs. In these stars, the force of gravity is balanced by the pressure of degenerate electron gas, this pressure being independent of temperature.

We also saw in Sect. 3.3.1 that stars with masses larger than about 8 $M_\odot$ go through several more nuclear burning cycles till the entire material in their core gets converted into iron. No further fusion reactions can take place inside these stars and the absence of energy generation leads to the collapse of the star. Under the high pressure, iron nuclei break up into their constituent protons and neutrons. The protons and electrons combine to form neutrons. All this results in a supernova explosion in which the outer layers of the star are completely blown away.

The remnant core is almost entirely made up of free neutrons and has only a tiny fraction of protons and electrons. Both protons and neutrons have spin half (the same as that of an electron) and Pauli's exclusion principle applies to them (see Box 3.1). Due to the high densities in the remnant core all three types of particle gases, i.e., electron gas, proton gas and neutron gas are degenerate, i.e., their lower energy levels are full and the particles are forced to occupy higher energy states.

At normal densities, any free neutron is unstable and decays in about 12 minutes to a proton, electron and an anti-electron neutrino ($\bar{\nu}_e$) as shown below.

$$\mathrm{n} \rightarrow \mathrm{p}^+ + \mathrm{e}^- + \bar{\nu}_\mathrm{e}$$

But this process, called beta decay, is suppressed in the very dense core of the star, as the electrons there are degenerate and the new electron produced in beta decay would have to have extremely high energy to be accommodated in the vacant higher energy levels of the electron gas. The energies of the emitted electrons can not be very large as the mass of the neutron is only a little larger than that of a proton, and as a result neutrons are prevented from decaying. The pressure of the degenerate neutron gas is independent of

temperature, just like the pressure of degenerate electron gas in white dwarfs. If these stars achieve hydrostatic equilibrium, i.e., if the pressure of degenerate neutron gas balances gravity, then it lasts indefinitely as the pressure of the degenerate neutron gas does not depend on its temperature and even if the remnant core loses its heat energy to the surroundings, the gas pressure does not decrease and keeps balancing gravity for ever. Such stars are called neutron stars. This is the second end stage in the life of stars. Neutron stars have sizes of about 10 to 20 km and densities about 3–6 $\times$ 10^{17} kg m^{-3}.

The existence of stars made up of neutrons was first proposed by Walter Baade and Fritz Zwicky only two years after the discovery of neutrons in 1932. The first theoretical model for neutron stars was constructed by J. R. Oppenheimer and D. M. Volkoff in 1939. Using earlier work of Tolman, they obtained an upper limit on the mass of a neutron star. It is known as the Tolman–Oppenheimer–Volkoff limit or the TOV limit in short. This is similar to the upper limit that Chandrasekhar had obtained for white dwarfs. That upper limit arose because of the fact that the pressure of degenerate electron gas can not balance the force of gravity if the mass of a white dwarf exceeded about 1.44 $M_\odot$. The upper limit on the mass of a neutron star arises because of the fact that the pressure of degenerate neutron gas can not balance the gravitational force if the mass of the star is larger than a certain value. This limiting value has been improved upon using more realistic models and lies between 1.5 and 3 $M_\odot$. The main reason for the large uncertainty in the value of the upper limit is the lack of accurate knowledge about nuclear forces. Stars having original masses larger than 8 $M_\odot$ but smaller than 25 $M_\odot$ throw away sufficient outer mass during supernova explosion, and end their lives as neutron stars.

Several properties of the neutron stars have been calculated. One important property is their rotation. Most stars in the Universe rotate around an axis passing through their centres. The speed of rotation rapidly increases, i.e., the period of rotation rapidly decreases, during the contraction of a star due to a principle called the principle of conservation of angular momentum. This is the same principle which allows an ice skater to spin faster on bringing her hands closer to her body. Looking at the typical rotation periods and sizes of stars and applying this principle, one can estimate that the rotation period of neutron stars can be of the order of a few milliseconds (ms).

Another property of neutron stars that was predicted was their magnetic fields. Most stars have magnetic fields. For example, Sunspots occur due to the presence of regions having higher magnetic fields. Similar to the conservation of angular momentum, there is conservation of magnetic flux. The magnetic flux through a given area is equal to the product of the value of the average

magnetic field there and the area. The principle tells us that the magnetic field of a star will get enhanced on contraction as its surface area decreases. A simple calculation tells us that the magnetic fields on the surfaces of neutron stars should be about 10^{14} times as strong as the average field on the surface of the Sun and about 10^{8} times stronger than the strongest magnetic fields created in any laboratory till date.

The temperature of a neutron star at the time of its formation could be estimated to be about 10^{9} K and is expected to drop to a few million degrees in a few hundred years. Such a neutron star, emitting its thermal energy as a black body (see Box 2.3) would emit mostly X-rays.

4.3.2 Award Winning Work: Discovery of Pulsars

On joining Ryle's laboratory in Cambridge, Hewish started working on scintillation. We are familiar with the phenomenon of scintillation which is responsible for the twinkling of the stars. It occurs due to the passage of star light through an irregular transparent media, i.e., the Earth's atmosphere. The conditions in the atmosphere are constantly changing which causes a continuous change in the density and refractive index at any point, thereby causing a change in the brightness and in the direction of the star perceived by us, i.e., its observed position on the sky. These, mainly the change in brightness, cause the appearance of twinkling or scintillation. Planets being much closer than the stars do not appear to be point sources like the stars and actually appear to be small discs. The change in the brightness and the position of the disc as a whole due to change in properties of atmospheric gases is not noticeable which explains why planets do not twinkle. Stars also do not twinkle when viewed through space telescopes which are above the Earth's atmosphere.

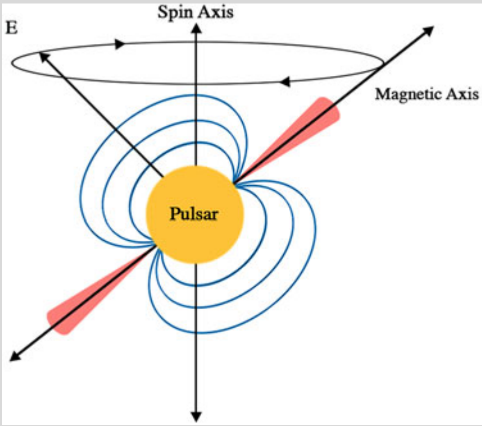
It turns out that radio sources also twinkle, and Hewish's first task at Cambridge was to understand this phenomenon. He showed, in 1951, that the scintillation of radio waves is caused during their passage through the ionosphere which is a layer in the upper part (48 to 965 km) of the Earth's atmosphere and contains ionized gases. Hewish showed that the scintillation occurred at a height of about 300 km and was able to measure the wind speeds there. In 1954 he observed scintillation in a radio source in the constellation of Taurus. After detailed study he concluded that the scintillation was caused by the irregularities in solar corona, which is the outermost layer in the atmosphere of the Sun, and extends up to several thousand km above the visible surface of the Sun (see Fig. 2.5). Hewish used scintillation to study the properties of gases in the corona, up to very large distances from the Sun. Later

he studied scintillation of extraterrestrial radio sources to study the properties of the interplanetary medium (i.e., the medium in between the planets) and went on to study the properties of solar wind, which is a constant wind of particles emanating from the surface of the Sun and has inhomogeneities in it giving rise to interplanetary scintillation. For scintillation to occur, the source has to have small angular size or angular diameter, i.e., the angle subtended by a diameter of the disc at our eyes. Thus, the presence or absence of scintillation provides limits on the angular sizes of radio sources. In 1965 Hewish found a scintillating radio source in the Crab nebula and suggested that to be the remains of the exploded star.

In 1965 Hewish undertook a large scale survey of 1000 radio galaxies, i.e., galaxies which emit radio waves, using scintillation as a tool. For this he needed to build a radio telescope having 2084 antennas. He did this together with his graduate student Jocelyn Bell. The task of setting up the telescope and testing it was completed in July 1967 and they undertook repeated observations of the radio sources, each source observed at different times of the day and night. On 6th of August 1967, Jocelyn noticed a fluctuating signal in a direction opposite to that of the Sun. As scintillation rarely occurs in the direction opposite to the Sun, she and Hewish first thought it to be due to some interference. But the source appeared again and again, each day coming 4 minutes earlier than the previous day. By November they were certain that the source was emitting pulses at intervals of about a second. As no such pulse-emitting cosmic source was known till then, they even considered the possibility of the signals being sent by some extraterrestrial beings. They ruled out various sources of human interference, waves reflected from the Moon's surface, television signals, orbiting satellites etc. They continued to search for similar sources elsewhere in the sky and by February 1968, they were confident of having found three more similar sources. These were named pulsars, short for pulsing stars. Initially, they thought the pulses were caused by pulsations of entire stars, but the period of the pulses was too small for them to be generated even by white dwarfs. In 1968, T. Gold suggested that the signal could be produced by rotating neutron stars and showed that the period should increase with time because of loss of energy in the form of radiation by the neutron star. Soon a pulsar was discovered in the Crab nebula with properties described by Gold, confirming the pulsar to be a neutron star, the stellar remnant of a supernova explosion as had been predicted earlier. Neutron stars have provided an opportunity for studying the behaviour of matter at extremely high densities, comparable or even larger than the densities inside atomic nuclei. Hewish received the Nobel Prize for the year 1974.

Box 4.3 Pulsar

The magnetic field of a neutron star is like that of a bar magnet and hence, the star has a magnetic axis, i.e., line joining the north and south poles of the magnet. The spin axis, that is the axis around which the neutron star spins is often not aligned with the magnetic axis. The rotation of the neutron star therefore leads to an off-axis rotation of the magnetic field associated with it. This produces electric field and its interaction with charged particles in turn leads to the emission of electromagnetic radiation. A very interesting fact is that the emitted energy is channeled into two narrow cone-shaped beams aligned along the magnetic axis as shown in the figure below. As the neutron star rotates, the beams sweep the sky along circles as shown in the figure, similar to the beam of a light house. If the Earth happens to be in a direction towards some point along the circle, e.g., towards E, it will receive a pulse of radiation once every rotation when the beam passes through the line of sight to an observer on the Earth. Over a length of time the observer will see a series of pulses, with the time between successive pulses being equal to the rotation period of the star. Some of the radiation is emitted at radio wavelengths, where the pulses are easiest to observe using a radio telescope. Pulses are also seen in the optical, X-ray, and other regions of the spectrum. As the rotation of a star can not be easily disturbed, pulsars provide very robust clocks.



4.4 Nobel Prize 2002: R. Giacconi

The Nobel Prize for the year 2002 was given to Riccardo Giacconi "for pioneering contributions to astrophysics, which have led to the discovery of cosmic X-ray sources." He received half the prize amount; the other half was shared by Raymond Davis Jr. and Masatoshi Koshiba.

Credit: http://www.nationalmedals.org/2003photos/giacconi/20050314_RKM_Medals_9316.JPG, Public Domain, https://commons.wikimedia.org/w/index.php?curid=2143158

Riccardo Giacconi was born in Italy in 1931. He received a Bachelor's degree from the University of Milan and continued there to obtain his Ph.D. in 1954. In 1956 he joined Indiana University, USA for research. In 1959, he joined the American Science and Engineering Inc., a small private research corporation in Massachusetts, to initiate a space science program for the corporation. There he was involved in exciting research, including development of a satellite and payloads for rockets, satellites, and aircraft. In 1973 he joined the Harvard–Smithsonian Center for Astrophysics. He was the founding director of the Space Telescope Science Institute in Baltimore from 1981 to 1993, which was set up as the science operation centre for the Hubble space telescope. He later worked as the director of the European Southern Observatory from 1993 to 1999. From 1999 to 2004, he was president of Associated Universities, Incorporated, which operates the National Radio Astronomy Observatory, USA. He passed away in 2018.

4.4.1 Background

So far we have discussed optical and radio telescopes. These are used to study the visible and radio radiation coming from astronomical sources. We know that electromagnetic radiation has been divided into 7 different bands (see Box 2.1). What about the remaining 5 types of electromagnetic radiations? Do astronomical sources emit them and can we detect them? The answer to both these questions is yes. However, these radiations from cosmic sources were not detected till about 70 years back. The main reason was that the Earth's atmosphere absorbs all radiation except the visible and radio which

are absorbed to a much lesser extent. So radiations of the remaining 5 types coming from astronomical sources, are not able to reach the Earth's surface and for detecting them we have to take our instruments above the Earth's atmosphere.

The interest in X-ray astronomy started when the Naval Research Laboratory in the USA began studying the structure of the Earth's upper atmospheric layer, namely the ionosphere. This study was important for understanding and the use of shortwave radio communications. For this study, instruments were flown aboard rockets to detect any X-rays coming from extraterrestrial sources. This resulted in the first detection of X-ray photons coming from the Sun in 1949. This was the first X-ray observation of a cosmic object and started the field of X-ray astronomy.

4.4.2 Award Winning Work: Discovery of Cosmic X-Ray Sources

Giacconi began his work in X-ray astronomy in 1959, about a decade after astronomers had first detected X-rays from the Sun. Giacconi conducted a number of observations with the help of X-ray detectors kept aboard rockets which flew to heights beyond much of the Earth's atmosphere. In 1962, he, along with his collaborators constructed an X-ray detector which was 10 times as sensitive as any other detector flown till then. This led to the first ever detection of X-rays from a source outside the solar system. It was named Scorpius X-1, which meant it was the first X-ray source detected in the constellation of Scorpius. This source turned out to have very unusual properties. Its X-ray luminosity, i.e., the amount of energy emitted per second in the form of X-rays, was a thousand times its luminosity in visible light. In comparison, the X-ray luminosity of the Sun is a million times smaller than its luminosity in visible light. Giacconi's team also detected several other cosmic sources of X-rays including the Crab nebula.

Giacconi's discovery kindled interest in X-ray astronomy among astronomers and several of them undertook rocket observations. These studies were limited by the short duration, of the order of a few minutes, of rocket flights above the atmosphere. To overcome this problem, in 1963 Giacconi along with H. Gursky, submitted a proposal to launch an X-ray telescope (see Box 4.5) with X-ray detectors on board a scanning satellite (scanning the entire sky). Till then, no X-ray telescope had been constructed and the only way to measure X-ray flux was to use X-ray detectors which could only measure the number of X-ray photons entering them. Thus, only strong X-ray emitters

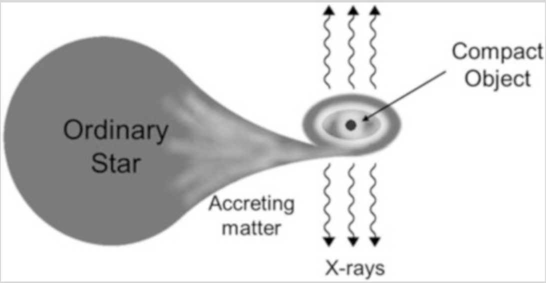
could be detected. An X-ray telescope would gather many more photons together which would then enter the detector, enhancing the sensitivity of the observations. Giacconi's proposal was accepted and the satellite UHURU (the Swahili word for freedom) was launched in 1970 from Kenya. However it only had detectors on board and not a telescope as was proposed by Giacconi, as it would take much longer to actually develop the concept of an X-ray telescope and to manufacture it. Having detectors on board a satellite increased the possible observation time from a few minutes allowed by the rocket flights, to a few years, i.e., by 5 orders of magnitude. The satellite could scan the entire sky in about 3 months. This increased the number of observed X-ray sources rapidly and it reached triple digits. Several of the sources could be identified with their optical or radio counterparts, which enabled their distances to be estimated and their X-ray luminosities to be measured. The identified sources included binary sources, supernovas, galaxies, quasars (about which we will learn in Chap. 6) and clusters of galaxies.

Several of the X-ray sources were pulsating sources, with periods of a few seconds while some others showed erratic pulsations with changes over tenths of a second. Detailed studies led to the conclusion that these sources are compact objects i.e., they could be white dwarfs, neutron stars, or black holes, in a binary system. A black hole is the third end stage in the life of stars, about which we will learn in Chap. 6. Here it will suffice to know that stars with initial masses larger than about 25 $M_\odot$ end up as black holes. The masses of their stellar remnants after they undergo supernova explosion are larger than the Chandrasekhar and TOV limits and they can not find equilibrium as white dwarfs or neutron stars. The other component of the binary is a normal star (see Box 4.4). The binary nature of these sources allowed the determination of the masses of the component (see Box 3.2). The masses of several of the compact components were found to be smaller than the Chandrasekhar and/or TOV limits which indicated that these objects could be white dwarfs or neutron stars. However, a few of the compact objects appeared to have masses between 3 and 6 $M_\odot$, and were above the Chandrasekhar and TOV limits. Thus they could not be white dwarfs or neutron stars. These then have to be black holes. The discovery of binary X-ray sources, i.e., systems having one compact component, provided several new insights which included (i) existence of binary stellar systems containing a neutron star or a black hole (ii) existence of stellar mass black holes (iii) measurement of mass, radius, rotation, and relation between pressure and density inside neutron stars (iv) a new source of energy due to gravitational infall (i.e., conversion of gravitational potential energy into thermal energy (see Box 4.4) which per nucleon is 100 times more efficient than nuclear fusion) and (v) a model for the energy

generation in quasars (as will be discussed in Chap. 6). It was also discovered, using X-ray observations, that there exists very hot gas between galaxies in clusters, which emits X-rays, and the mass of this gas is as much as the total mass of all galaxies in the cluster.

Box 4.4 Binary X-Ray Sources

An X-ray binary is a system of binary stars out of which one star is an ordinary star burning nuclear fuel, while the other star is a compact object i.e., a white dwarf, neutron star, or black hole. During the course of evolution, the ordinary star expands and the material in its outermost layer comes sufficiently close to the compact object to be influenced by its gravity and is pulled by it. This accreting matter, i.e., the matter falling on the compact object, forms a disc called the accretion disc, around the compact star before actually falling on to it (see figure below). The infalling matter gets heated due to the decrease in its gravitational potential energy and its conversion to thermal energy. The temperature of this accreting matter reaches up to about a million K at the place it falls, causing hot spots in localized regions. At that temperature the black body emission by the matter peaks in X-rays. These X-ray sources are periodic. The periodicity in the emission of X-rays is due to the rotation of the compact star.

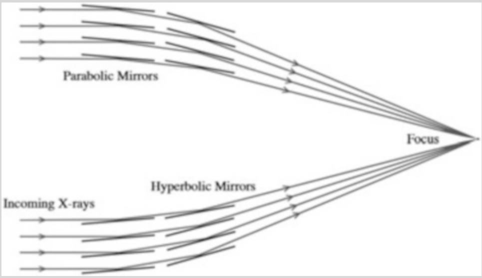


Giacconi worked out the operating principles for a telescope that could focus X-rays and produce images of the source (see Box 4.5). He constructed the first X-ray telescope, which was placed in another X-ray satellite observatory which was named the Einstein Observatory and launched in 1979. This increased the sensitivity of detection of X-ray sources by a factor of a thousand which made it possible to study stellar atmospheres and supernova remnants. It identified many more X-ray binary systems, and also detected X-ray sources in other galaxies. In 1976, Giacconi proposed an even more powerful instrument, which was finally launched in 1999 and was named the Chandra X-Ray Observatory after Nobel laureate S. Chandrasekhar. Giacconi

was awarded the Nobel Prize for his enormous contributions to the field of X-ray astronomy in particular and to astronomy in general.

Box 4.5 X-Ray Telescopes

We have seen earlier that a reflecting telescope has a mirror which reflects the rays falling on it and collects them at its focus. We know that X-rays have great penetrating power and pass through most substances. How then can we have a mirror which will reflect X-rays and focus them at one point? Well, this is possible if the X-rays are made to fall on a mirror at a grazing incidence, i.e., if the incident X-rays are almost parallel to the plane of the mirror as shown in the following figure. However, as the plane of the mirror is almost parallel to the incoming X-rays, the area of the mirror as seen by the X-rays is extremely small and very small amounts of X-rays can be collected by the mirror. To increase the effective X-ray collecting area, multiple layers of mirrors are used as shown in the figure.



5

Extrasolar Planets

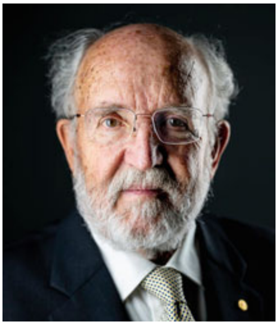
5.1 Introduction

The Nobel Prize in the category of extrasolar planets or exoplanets was given in the year 2019 to two astronomers: Michel Mayor and Didier Queloz. Their work concerns planets around stars other than the Sun and is purely experimental. It started in the early nineteen nineties and the award winning work was announced in 1995. For the most part they worked together.

5.2 Nobel Prize 2019: M. Mayor and D. Queloz

The Nobel Prize for the year 2019 was given to Michel Mayor and Didier Queloz "for discovery of an exoplanet orbiting a solar-type star." They shared half of the prize amount; the other half was awarded that year to cosmologist James Peebles.

P. Khare, *Nobel Prizes in Astronomy*, Astronomers' Universe,
https://doi.org/10.1007/978-3-031-29639-0_5



Credit: Nobel Media, photo: A. Mahmoud

Michel Mayor was born in 1942 in Switzerland. After his post graduation from the University of Lausanne in 1966, he obtained his Ph.D. from the Geneva Observatory in 1971. From 1971 to 1984 he worked as a research assistant at the Geneva Observatory, after which he joined as a faculty in the University of Geneva. He was the director of the Geneva Observatory from 1998 to 2004. He retired in 2007, but continues as Professor Emeritus since then.

Didier Queloz was born in 1966 in Switzerland. After finishing his Master's degree from the University of Geneva, he went on to obtain his Ph.D. in 1995 under the guidance of Mayor. He was a postdoctoral fellow at the University of Geneva from 1995 to 1997, after which he was a visiting fellow at the Jet Propulsion Laboratory, Pasadena. In 2000, he returned to Geneva and became a professor at the University of Geneva in 2008. In 2013 he was also appointed a professor at the Cavendish laboratory under Cambridge University. There he established the Cambridge Exoplanet Research Centre.

5.2.1 Background

For ages, mankind has pondered over the question "Are we alone in the Universe?" The possibility that advanced life exists elsewhere in the vast Universe, and may be trying to contact us by sending messages, has also been considered since long. Reports of Unidentified Flying Objects (UFOs) keep appearing in the press from time to time. The possibility that these are spacecrafts sent by aliens is taken seriously by a few.

It has long been understood that the extraterrestrial beings may be located far away and may be too sparsely distributed in the Universe to physically go and look for them. So the only way to discover or look for them is through the exchange of messages. Radio waves are the best way to send messages over long distances as these waves do not get appreciably absorbed or scattered by the dust and gas that they encounter along the way. If indeed a technologically advanced civilization exists anywhere in the Universe, it also would try to communicate with other civilizations by sending out messages via radio waves. A project to receive such messages sent by extraterrestrial beings who may be present anywhere in the Universe has been in operation since 1984 in Mountain View, California. This project is run by a non-profit research organization and is called SETI, short for "Search for Extra Terrestrial Intelligence" and has about 100 scientists working on the research program employing radio telescopes which are located in the Cascade Mountains of California. However, so far no meaningful signal has been picked up by the SETI scientists.

Over the past few centuries, we have gathered a lot of information about the contents of the Universe. We know that galaxies are the building blocks of the Universe. Galaxies are vast collections of stars and rarefied gaseous and dusty interstellar matter. Some rarefied gaseous material also exists in the space between galaxies and is called intergalactic matter. Radiation and some elementary particles are present throughout the Universe. If any life exists in this vast Universe, in the form that we can recognize, i.e., which is basically similar to life on the Earth, it can only exist on planets. Stars are too hot and interstellar and intergalactic matter is either too hot or too cold, and also too rarefied for such life to thrive there. Till about 1992 we only knew of planets in the solar system, and as far as we could tell, there was no sign of any life, let alone intelligent life, existing on the solar system planets other than the Earth. The solar system planets are either too hot, like Mercury and Venus, or too cold and gaseous, like Jupiter, Saturn, Uranus or Neptune. So if we want to search for extraterrestrial life, it has to be on extrasolar planets.

Assuming there are planets around at least some of the other stars in the Milky Way, can we see them directly through telescopes or get their images, like we can do for the planets in the solar system? The answer is that, though not impossible, it is very difficult to do so. This method works only if the planet is sufficiently far from the glare of the star and is emitting infrared radiation. The reason why direct imaging is difficult is that, generally, planets do not emit their own light, i.e., they are not primary sources of light. They only reflect the light of their star, i.e., the star around which they are revolving, that falls on them. As planets are very small in size as compared to the star and also are at a distance from their stars, the amount of star light falling on them is only a small fraction of the total amount of light emitted by their star and out of this, only a small fraction is reflected. For example, Jupiter reflects about one billionth of the luminosity of the Sun. A small portion of the reflected light falls on the Earth. Even in our solar system, we see that the planets are much fainter than the Sun. So planets around other stars will be too faint to be observed directly. Very few, around 25 of the total number of close to 5000 planets observed so far, have been through direct imaging. How then, do we detect these planets? There are two main methods to detect extrasolar planets: the radial velocity method and the transit method. Radial velocity of an object is its velocity along the line of sight to the object, i.e., the line joining the observer and the object. More than 90% of the extrasolar planets have been discovered using the above two methods. Let us look at these methods.

The Radial Velocity Method

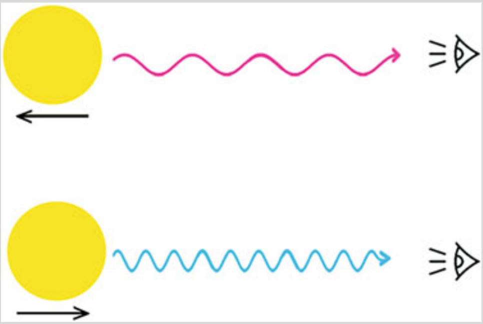
When two astronomical bodies revolve around each other like two stars in a binary system, or a planet and its star, their orbits are ellipses with their centre of mass at the common focus (Box 3.2). For the case of a star and its planet, usually, the two masses are very different and the centre of mass is very close to the star. For example for the Earth-Sun pair, the centre of mass is in fact, inside the Sun and is very close to its centre. The rotation of the Sun around the centre of mass in this case, is only a slight wobble, while the Earth goes around it. The centre of mass of the Sun-Jupiter system lies just above the solar surface. The centre of mass of the solar system as a whole keeps changing its position because the planets keep changing their position all the time.

In the radial velocity method an extrasolar planet is detected by observing the radial motion of its star with respect to the observer, caused by the planet. The radial motion can be detected through an accurate measurement of the radial velocity of the planet using the Doppler effect (see Box 5.1),

which is a wave phenomenon and applies to all types of waves like sound waves, electromagnetic waves etc. As we have seen, a wave has a characteristic wavelength. For example, the wavelength of visible light decides its colour. The Doppler effect tells us that the wavelength of a wave emitted by a source is different from the wavelength measured by an observer if the two are moving towards or away from each other, i.e., have a nonzero radial velocity with respect to each other. The difference in the two wavelengths depends on the value of the radial velocity between the two.

Box 5.1 Doppler Effect

The Doppler effect was discovered by Austrian physicist, Christian Doppler in 1842. When a source of a wave is stationary with respect to the observer, the wavelength received/measured by the observer is identical to the wavelength of the wave emitted by the source. However, if the source is moving towards or away from an observer, i.e., has nonzero radial velocity then the wavelength measured by the observer is different from the value of the emitted wavelength. If the source is moving towards the observer, the measured wavelength is smaller than the emitted wavelength, and it is larger when the source is moving away from the observer. We have all witnessed this effect for sound waves. The sound of the siren of a police car or ambulance changes pitch as it passes us. The wavelength is smaller, i.e., the pitch is higher when the source is approaching us and it is lower after the source has passed us. It does not have to be the source which has to move. Even if the source is stationary and the observer is moving, the Doppler effect is observed. It is the relative radial velocity which is responsible for the Doppler effect.



This effect has also been observed for light. If a source of light emitting yellow light say, as shown in the figure, is moving towards us (or we are moving towards the source) the wavelength of the light received by us, i.e., as measured by us, appears to be smaller than that of the emitted light, i.e., the wavelength shifts towards that of the blue colour. This change in wavelength is called blueshift. On

(continued)

Box 5.1 (continued)

the other hand, if the source is moving away from us, the wavelength appears to us to be larger than its emitted value, i.e., it shifts towards that of red colour, and the change is called redshift. Note that the effect has been exaggerated for the purpose of illustration. In practice it will require relativistic velocities for such change in colour to occur. Quantitatively, redshift is defined as the fractional change in wavelength, i.e., the ratio of the change in wavelength to the original or emitted wavelength. For small radial velocities, this fraction is equal to the ratio of the radial velocity of the source divided by the velocity of light in vacuum. For velocities close to the velocity of light the relation between redshift and velocity is somewhat more complex.

Thus, if we know the wavelength of the radiation emitted by the source and we can measure the wavelength of the radiation received by us, we can determine the radial velocity of the source. This method can be applied to measure the radial velocity of a star. However, a star does not emit radiation of one particular wavelength, but emits a continuous black body spectrum (see Box 2.3). Due to the radial motion of the star, the whole continuous spectrum shifts in wavelength. It is not possible to determine the redshift or blue shift of a continuous spectrum. However, we are helped by the fact that stellar spectra have black lines superimposed on the continuum, as can be seen in the solar spectrum shown in Fig. 5.1. This has been obtained by passing sunlight through an instrument called high (spectral) resolution spectrograph which spreads out light of different wavelengths. The higher the spectral resolution of the spectrograph, the larger is the spread of light over wavelengths.

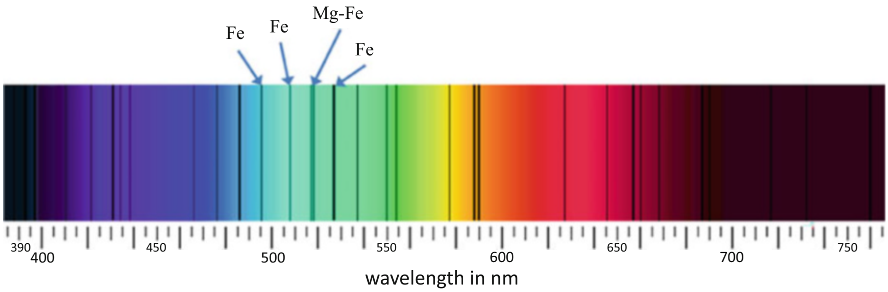


Fig. 5.1 High resolution spectrum of the Sun obtained by passing sunlight through a spectrograph. Dark lines are produced due to absorption of radiation by atoms and ions present in the solar atmosphere. The elements producing some of these are labeled in the figure

The dark lines, called absorption lines, are produced due to the absorption of light at specific wavelengths by the ions, atoms, and molecules present in the cooler atmosphere of a star, i.e., the layer outside its photosphere (see Sect. 2.3.1 and Fig. 2.5). The specific wavelengths at which absorption takes place are decided by the structure of the atoms of the elements present in the stellar atmosphere and can be considered to be the fingerprints of these elements. Some of the lines produced by iron and magnesium are shown in the figure. The wavelengths of most of these lines have been determined accurately in the laboratory when the source is at rest with respect to the observer. It is easy to identify the lines in the stellar spectra, as even though the wavelengths of the lines may be different from their laboratory values due to the Doppler effect, the changes in wavelengths of different lines are such that the ratios of their wavelengths remain unchanged. The ratios of the observed wavelengths of absorption lines in a stellar spectrum thus allow us to identify them and to measure the radial velocity of the star with respect to the Earth.

Note that if the Earth happens to lie in a direction which is perpendicular to the plane of orbit of a planet around its star, the star will not have any radial velocity with respect to the Earth as its velocity will always be in a plane perpendicular to the line of sight. Using the radial velocity method, we can only hope to detect those planets whose plane of rotation around their respective stars are inclined at an angle smaller than 90° with respect to the line of sight from the Earth. The radial velocity will be largest when this angle is near zero, i.e., when the line of sight from the Earth is in the same plane as the orbit of the planet.

If a star is found to have radial velocity which changes periodically, it indicates the presence of a planet orbiting around it. The star will also have periodic velocity in case it is a binary star and is orbiting around another star. However, the velocities of stars in binary systems and those caused by a planet are very different in magnitudes. For example, the velocity of the Sun caused by the influence of the Earth is about a few cm s^{-1} and that caused by Jupiter is of the order of a few m s^{-1}. On the other hand, the typical velocities of stars in binary systems are of the order of few km s^{-1}. Thus, very accurate measurements of radial velocities are needed for detection of extrasolar planets. The magnitude of the shift of the absorption lines along with the mass of the star (determined through other methods) helps us to estimate the mass of the planet using Kepler's laws, and the periodicity of the shift gives the rotation period of the planet.

Transit Method

We all know that a solar eclipse occurs when the Moon comes in between the Earth and the Sun and its shadow is cast on the Earth. Persons in the shadow can not see the Sun (or a part of it) for a short while during which the Moon is passing in front of the Sun. The eclipse can be total, partial, or annular depending on the distances and geometry involved. Similar to the Moon, other inner planets namely, Venus and Mercury also cause solar eclipses when they happen to come in between the Earth and the Sun. These planets are much farther away from the Earth as compared to the Moon, and they cover only a tiny circular portion of the solar disc. Thus, these events do not appear like eclipses and the planets appear like tiny dark circular dots transiting across the solar disc. The events are therefore called transits. The same phenomenon takes place for an extrasolar planet-star system as viewed from the Earth, when the line of sight to the planet is close to the plane of its orbit.

During a transit, the brightness of the star, which is defined as the amount of energy received per unit area per unit time by an observer, decreases by a small amount as the planet obstructs some of the starlight from reaching the observer. The dip in stellar brightness depends on the size of the planet and its distance from the star. This dip is seen periodically, once during every orbit of the planet. A sketch of the dip in the brightness of a star due to planetary transit is shown in Fig. 5.2. The dip is highly exaggerated in the figure for purpose of illustration. In reality it is extremely small and highly sensitive instruments are needed to measure it.

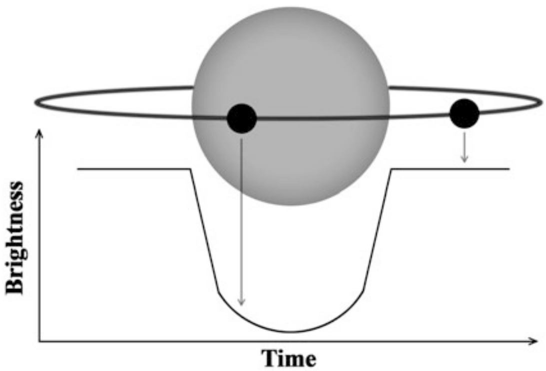


Fig. 5.2 The upper part shows two positions of a transiting planet around its star and the bottom part shows the change in the brightness of the star due to the transit of the planet across the star. The decrease in brightness has been greatly amplified for the sake of illustration

5.2.2 Award Winning Work: Discovery of an Exoplanet Orbiting a Solar Type Star

During his early years as a research associate at the Geneva observatory, Mayor developed an instrument for accurate measurement of radial velocities of stars. Till then, photographic plates were used to record stellar spectra and a painstaking study of these plates had to be undertaken to measure the Doppler shifts of the spectral lines therein. Mayor developed an electronic spectrometer, which he named CORAVEL (COrelation-RAdial-VELocities). The spectrometer produced a high resolution spectrum of the source which was projected on a glass plate coated with chromium, except at places where atomic absorption lines were expected. Thus, the plate was completely opaque except at the laboratory wavelengths of absorption lines. If the radial velocity of the source was zero, i.e., if the source was stationary with respect to us, the absorption lines will fall exactly at the positions of the transparent lines on the plate. In this case, the light transmitted by the glass plate would be minimal as only the residual light in the absorption lines, which are dark, would be able to pass through the plate. If the radial velocity was not zero, the positions of the absorption lines in the spectrum would not match those on the glass plate and the amount of transmitted light would be larger. The device would then allow the measurement of how much the observed spectrum had to be shifted to get the minimum transmitted light, thereby giving the magnitude of the Doppler shift and thus, that of the radial velocity of the star. Thousands of absorption lines were used in this process, which increases the accuracy of measurement.

In 1977, Mayor could measure radial velocities with an accuracy of 300 m s^{-1}. It was about 4000 times more efficient than photographic plates. The instrument was used to study dynamics of old stellar clusters called globular clusters and pulsation of particular type of stars called the Cepheid variables about which we will learn in Chap. 8. He then used his instrument to measure the radial velocities of the stars which were thought to be binary stars. In 1991 his study showed that a subset of these stars may not be in binary stellar systems, but may be stars having companions with substellar masses and so could be planets. Further studies of these stars required higher accuracy in measurement of radial velocities than what the spectrometer CORAVEL was capable of, as the radial velocities caused by substellar objects are extremely small as we have seen above.

After Queloz joined Mayor as a graduate student, they worked on improving the capabilities of CORAVEL, as new technology in the form of electronic detectors and fiber optics had become available then. They developed a

spectrometer, which they named ELODIE which made use of both these advances in technology. This spectrometer could measure velocities with the accuracy of 10–15 m s^{-1}, which was a factor of 20 to 30 better than CORAVEL. In 1994, Mayor and Queloz started a survey at the Geneva observatory, of 142 Sun-like stars which were suspected to have substellar components. For one particular star, 51 Pegasi, they could rule out other possible explanations for the observed radial velocity and conclusively showed that the star had a planet orbiting around it which was causing the radial velocity of the star to change by 59 m s^{-1}.

The planet was named 51 Peg b. It orbits 51 Pegasi every 4.23 days. This was the first ever exoplanet which was discovered to be orbiting a Sun-like, hydrogen burning star. It was also the first planet whose spectrum was observed. Later studies showed it to be a hot, Jupiter like planet. It is about half the size of Jupiter and is not thought to be habitable. According to the theory of planet formation, giant gaseous planets could only form at large distances from their stars and would have orbital period longer than 10 years. The discovery of 51 Peg b did not fit this picture. It is now believed that the planet indeed formed far away from the star but moved closer to it due to gravitational interaction with the material of the disc around the star (out of which planets form) in a process known as orbital migration of a planet. Although a few planets had been discovered before 51 Peg b, some were found around pulsars or dead stars and the observations for the others were not completely convincing. Mayor and Queloz's discovery showed for the first time, conclusive evidence for a planet orbiting a Sun-like star. They were awarded the Nobel Prize for their discovery.

Mayor and Queloz have continued to improve on their spectrometer. After ELODIE there have been more spectrometers, which along with the year of making and the accuracy of velocity measurement are: CORALIE (1998, 6 m s^{-1}), HARPS (2003, 1 m s^{-1}), and ESPRESSO (2018, 10 cm s^{-1}). HARPS was used on a bigger, 3.6 m diameter telescope at the ESO observatory at La Silla in Chile. Along with other astronomers, Mayor and Queloz undertook a survey for detecting exoplanets and in 2009 they identified the first extrasolar planet Gliese 581c which was in the star's habitable zone and may be conducive to the presence of life. With the HARPS spectrograph, they have now detected a new population of Earth to Neptune mass planets (planets with masses between 1 and 20 Earth-masses) which is a population of extremely common planets orbiting Sun-like stars. The search for more exoplanets continues with the aim of understanding the formation, chemical composition, and evolution of planets.

6

Black Holes

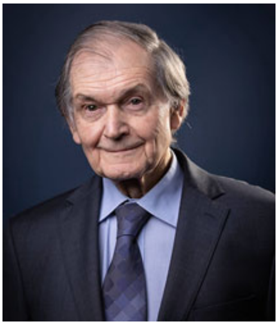
6.1 Introduction

The Nobel Prize in the category of black holes was awarded in the year 2020 to three astronomers: Reinhard Genzel, Andrea Mia Ghez, and Roger Penrose. This concerns different aspects of black holes. The work of Penrose is purely theoretical and highly mathematical. It deals with black hole formation. His award winning work was done in 1965. The work of Genzel and Ghez is observational which resulted in the discovery of a supermassive compact object, expected to be a black hole, at the centre of the Milky Way Galaxy. The work has been carried out over about a quarter of a century and the award winning works were published in 1996 and 1998.

6.2 Nobel Prize 2020: R. Penrose

The Nobel Prize for the year 2020 was given to Roger Penrose "for his discovery that black hole formation is a robust prediction of the general theory of relativity." He received half the prize amount; the other half was shared by Reinhard Genzel and Andrea M. Ghez.

P. Khare, *Nobel Prizes in Astronomy*, Astronomers' Universe,
https://doi.org/10.1007/978-3-031-29639-0_6



Credit: Nobel Prize Outreach, photo: Fergus Kennedy

Roger Penrose was born in 1931 in the United Kingdom. After graduating in mathematics from the University College London, he obtained Ph.D. in mathematics from St John's College, Cambridge in 1958. He spent two years (1959–1961) in the USA at Princeton University and Syracuse University on a NATO Research fellowship. After spending the next two years as a researcher at St John's College, he went back to the USA as a visiting professor at several universities. He was a faculty in the applied mathematics department at Birkbeck College, London from 1964 to 1973. From 1973 he held the Rouse Ball Chair of mathematics at Oxford University. He wrote several books including two popular ones on Consciousness, namely, "The Emperor's new mind" and "Shadow's of the Mind". He also made interesting contributions to the impossible diagrams made famous by the artist M. C. Escher. His creations "Penrose triangle" and "Penrose tilings" are well known.

6.2.1 Background

If the initial mass of a star is larger than about 25 $M_\odot$, it evolves in a way similar to the stars with masses roughly between 8 and 25 $M_\odot$ and, as described in Sect. 3.3.1 and in Sect. 4.3.1, undergoes a supernova explosion. But, for these stars, the mass of the stellar remnant made up of neutrons, left after the explosion, is larger than the TOV limit. It thus can not achieve hydrostatic equilibrium as the pressure of the degenerate neutron gas can not balance the huge gravitational force. As there are no other elementary particles with spin 1/2, which are abundant and can provide the necessary degenerate pressure to balance gravity, the remnant is doomed to collapse for ever. This is the fate of all stars with initial masses larger than about 25 $M_\odot$. Once the size of the star becomes smaller than a certain limiting value, it becomes what we call a black hole which is described below.

Black Holes

As we know, a ball thrown up from the surface of the Earth, travels up to a certain distance, its velocity decreasing with time because of the gravitational pull of the Earth. When the velocity becomes zero, the ball stops and then falls back on the Earth because of the same pull. If we throw the ball with higher velocity, it will be able to reach a greater height before falling back. What if we throw the ball with very high velocity so that it will reach infinite height, i.e., the pull of the Earth will not be sufficient to stop it? Yes, this is possible and in this case, the ball will continue to rise up, albeit with decreasing velocity, but will not fall back on the Earth. The minimum velocity with which we have to throw the ball (or any object) for this to happen is $11.2\,\mathrm{km\,s^{-1}}$. This value is called the escape velocity of the Earth, meaning, it is the minimum velocity with which an object has to be thrown vertically upwards from the Earth's surface for it to permanently escape the gravitational pull of the Earth.

The value of $11.2\,\mathrm{km\,s^{-1}}$ is obtained by using the mass and the radius of the Earth. Every astronomical object has its own escape velocity. For example, its value on the surface of the Moon is $2.4\,\mathrm{km\,s^{-1}}$, the Moon's gravity being weaker than that of the Earth, while that on the surface of the Sun is $618\,\mathrm{km\,s^{-1}}$. Obviously, the higher the mass, the higher is the escape velocity because of higher gravitational pull. However, the more important factor for us here is the size of the object. For two objects with the same mass, the smaller the size, the higher is the value of escape velocity. Thus, as a star contracts, its escape velocity increases.

As we saw above, the stellar remnants of supernova explosions of stars with initial masses larger than about $25\,M_\odot$ can not achieve a balance between pressure and gravitational force, and are doomed to contract for ever. In this process of contraction and consequently that of increasing escape velocity, a stage comes when the radius of the star becomes so small that the escape velocity becomes larger than the velocity of light in vacuum, i.e., 300,000 km $\mathrm{s^{-1}}$. Let us call this radius R_{BH}, which depends on the mass of the star.

We know from the special theory of relativity (see Box 2.5) that nothing can travel faster than the speed of light in vacuum, and therefore, neither light nor information conveyed in any other form can come out of an object having radius smaller than R_{BH}. Any light falling on the object cannot come out again, i.e., it cannot get reflected. Remember that we can see an object because of the light it emits, scatters or reflects. Thus, for all practical purposes, a star with its radius smaller than its R_{BH} will become invisible and all we could see at its position would be a black hole, which justifies the name given to these stellar remnants. These black holes are predicted by Newton's theory of gravity

(combined with special relativity), and are called Newtonian black holes. The value of R_{BH} for a star with a mass of 25 $M_{\odot}$ is 72.5 km, while for the Sun it is about 3 km.

General Theory of Relativity

Till the early twentieth century, Newton's theory of gravity was a very successful theory. It could explain the motion of planets, motion of projectiles, occurrence of tides, periodicity of comets, shapes of galaxies, etc. It only had a couple of small problems. One was that the gravitational force acted between two bodies even without any medium present between them. Newton himself was deeply unhappy about this aspect. The second problem was that, the force was felt instantaneously, e.g., if one body was somehow removed, the other body would instantaneously feel the absence of its gravitational attraction. After Einstein gave his special theory of relativity according to which no object or information can travel faster than the velocity of light in vacuum, this problem became even more serious, as the gravitational force was felt instantaneously in Newton's theory, and so the information about the presence of another body seemed to travel with infinite speed which was inconsistent with the special theory.

Apart from these two issues, there also was a small difference between the prediction of Newton's theory for the orbit of Mercury and observations. The orbit of Mercury does not remain fixed in space. It would have been so if it was the only planet in the solar system. The gravitational force due to the rest of the planets makes the orbit precess around the Sun, i.e., the semimajor axis of the orbit of Mercury rotates around the Sun. This rate of precession, calculated on the basis of Newton's theory, is 532 arc sec per century. Its observed value was 575 arc sec per century. The difference, though tiny, is present, and could not be accounted for using Newton's theory.

In 1915 Einstein came up with his theory of gravity which is called the General Theory of relativity. The theory had very novel ideas, required complex mathematics, and was initially very difficult to understand. According to this theory, gravitational field induces curvature in space-time. In absence of any gravitational field, space-time is flat. A massive object bends the space around it just like a heavy ball placed on a rubber mattress depresses the originally flat surface of the mattress. It should be noted that the surface of the rubber mattress is two dimensional while space-time is 4 dimensional. This curvature in space-time affects the motion of objects. The higher the mass of the object producing the gravitational field, the higher will be the curvature

and thus, the higher will be the effect on the motion of other nearby bodies. Thus, even though there is no physical contact between the Sun and a planet, the mass of the Sun bends the space-time around it and the motion of the planet around the Sun is the effect of this curvature. The gravitational field which is perceived through the curvature does not travel instantaneously, but travels with the speed of light. This solved both problems faced by Newton's theory. The curvature due to any given mass distribution can be obtained by solving equations given by Einstein, which are known as Einstein's equations.

For weak gravitational fields and nonrelativistic motions, the general theory of relativity gives the same results as those of Newton's theory, thereby keeping the successes of Newton's theory intact. The results of the general theory of relativity differ from those of Newton's theory when the gravitational fields and velocities are high. This theory is essential for understanding the structure of very compact stars and the large scale structure of the Universe. It is also useful in day to day life, e.g., it is essential for accurate time keeping in the satellites used for the Global Positioning System (GPS). Using Newton's theory for this would result in large errors in the positional predictions of the GPS and render it useless.

A few important predictions of the general theory of relativity are as follows.

- Bending of light: The general theory predicts that the motion of a light ray is affected by gravity and light bends while traveling close to massive objects, being attracted by it. Before this theory, light rays were believed to travel in straight lines. That light rays do indeed bend while passing close to the Sun, was observationally verified by Eddington by measuring the apparent position of stars very close to the Sun. This observation could be made for a few stars during the total eclipse of 1919 which was seen in Africa. The apparent positions of these stars were different from their actual positions which were measured 6 months earlier when the stars were seen in a direction opposite to that of the Sun (due to the revolution of the Earth around it) and were visible in the night sky. The differences in the real and apparent positions of the stars were consistent with the predictions of the general theory of relativity. The bending of light also produces interesting effects like gravitational lensing which causes distorted or multiple images of a distant source lying behind a massive object, since gravity bends light like a glass lens does. Several instances of gravitational lensing have been observed.
- Precession of orbits: The general theory of relativity predicts the precession of closed orbits of objects moving around a massive object. The rate of

precession of the orbit of Mercury calculated using this theory exactly matches the excess precession that could not be explained by Newton's theory. Precession of orbit has also been observed in orbits of binary systems. Thus, Einstein's theory solved the third problem faced by Newton's theory.

- Gravitational redshift or Gravitational time dilation: We have seen earlier that a ball thrown up from the Earth's surface loses its energy and slows down due to the gravitational pull of the Earth. According to the general theory of relativity, the same effect takes place for light particles, i.e., photons as well. Photons emitted by a massive object lose energy while coming out of the object. The loss in energy can not decrease the velocity of the light rays as their speed is fixed according to the special theory of relativity. The loss in energy of photons results in a decrease in its frequency (remember that the energy of a light particle, i.e., photon, is directly proportional to its frequency, see Box 2.1) or increase in its wavelength. The gravitational pull thus causes redshift of the light emerging from a massive object. As this redshift is caused by gravity, it is called gravitational redshift. Light falling on a massive object will experience the opposite effect and the photon will gain energy due to the gravitational pull of the object and will get blueshifted. These effects were verified by experiments conducted by Pounds and Rebka in 1959. This effect can be looked at in another way. We know that frequency is the number of oscillations in one second. The frequency of a photon near a massive object is higher than what it is when it goes away from the object. Said differently, the number of oscillations per second when the photon is near a massive object is higher than the number of oscillations per second when it goes away from the object. Instead of seeing this as a change in the rate of oscillation, we can assume that the time interval of 1 second near the object is larger than what the interval is when the photon is away from the object. This can be stated as "the rate at which a clock ticks is affected by the gravitational field." A clock located at a place having a higher gravitational field (near a massive object) will run slower than a similar clock located at a place having a lower gravitational field (away from the object). This is known as gravitational time dilation.

Einstein formulated his equations in 1915. The first simplest exact solution to these equations was obtained by the German physicist, Karl Schwarzschild, in 1916 and is called the Schwarzschild solution. It describes the spherically symmetric gravitational field around a point object with a given mass. It showed that there is a particular value of the radius, later named the Schwarzschild radius, from within which light can not escape to the outside world. This radius is proportional to the mass of the object. Thus, once the

size of a star becomes smaller than its Schwarzschild radius, it will become what is known as a Schwarzschild black hole. The surface of the sphere of radius equal to the Schwarzschild radius is known as the event horizon of the star. No material particle or radiation, i.e., information in any form can escape outside from inside the event horizon. Note that the value of Schwarzschild radius is the same as R_{BH}, obtained for the radius of a Newtonian black hole.

It can be shown that the only properties that a black hole can have are mass, rotation or spin, and electric charge. However, as black holes form from stars and stars do not have charge, black holes are not expected to have any charge. A rotating black hole is known as a Kerr black hole, after Roy Kerr who discovered the corresponding solution of Einstein's equations in 1963. So a black hole is a pristinely simple object. Nevertheless, black holes are physically and mathematically extraordinarily interesting, and are the key to much of modern astrophysics.

As we can not see a black hole, how do we know if black holes really exist or not? Information about the presence of a black hole comes from the gravitational force it exerts on the surrounding objects. Observational evidence for the presence of stellar mass black holes comes through the observations of binary X-ray sources as discussed in Sect. 4.4.2. Direct observational evidence for the existence of black holes has been obtained in the last few years through the observation of gravitational waves emitted during coalescing binary black holes, as we will see in Chap. 7. In addition to the stellar mass black holes, there exist supermassive black holes with masses around a million solar masses. These exist at the centres of galaxies. The evidence for these is described in the next section and is the subject of another Nobel Prize.

6.2.2 Award Winning Work: Black Hole Formation in the General Theory of Relativity

In general relativity, gravitational field is the measure of the curvature of space-time. The higher the field, the larger is the curvature. Schwarzschild's solution, given in 1916, described the spherically symmetric gravitational field around a point object. According to this solution, the curvature of space-time grew without bound as one approached the point mass. At the centre, i.e., the position of the point mass, it became infinite and that point is called a singularity in mathematical terms. The Schwarzschild solution had another singularity as well. This was at all points on the surface of the sphere having radius equal to the Schwarzschild radius, i.e., on the event horizon. The surface had weird physical and mathematical properties. One was that time would

stand still at the event horizon due to the infinite gravitational time dilation there. These properties could not be understood at that time. However, for most real stars, the Schwarzschild radius is much smaller than the actual radius of the star and what happens when the radius of the star becomes smaller than the Schwarzschild radius was hypothetical and did not attract much attention.

As we have seen in Chap. 3, Chandrasekhar proved in the 1930s that (see Sect. 3.2.2) stars with final masses, i.e., the masses of the stellar remnants after having shed their outer layers in explosions leading to planetary nebula formation, larger than 1.44 $M_\odot$ can not achieve equilibrium as white dwarfs. They are doomed to undergo further gravitational collapse. This rekindled scientists' interest in the question of black holes and singularities. As seen earlier, Eddington ridiculed the idea of unhindered gravitational collapse of a star. Even Einstein did not believe in such a collapse and in 1939 stated: "That time would stand still at the Schwarzschild radius is a patently absurd conclusion and if true, an object falling into a black hole would hover at the Schwarzschild radius for ever." He believed that these effects are the artifacts of the idealization that the star is a point mass. Oppenheimer and Snyder obtained solutions of Einstein's equations for a collapsing star in 1939 and concluded that the star will continue to collapse to a radius smaller than its Schwarzschild radius. However, these results were also dismissed by most scientists as being due to certain assumptions made in the calculations. The discovery of the quasar (described in the next section) in 1963, stimulated renewed interest in the question of gravitational collapse.

The main assumption in all these calculations was of spherical symmetry. In reality, there could be perturbations which will disturb the spherical symmetry and it was not clear if such objects will collapse to a singular point. Thus, the fact that a star can indeed collapse to be inside its Schwarzschild radius was not accepted as proven in general relativity for a long time. It was also believed that singularities could be an artifact of the choice of the coordinate system. This can be understood through the example of the Earth and the coordinate system of longitude and latitude. Any point on the Earth can be described uniquely by its longitude and latitude, except for the two poles which do not have a unique value of the longitude. However, we know that there is nothing mysterious about the poles and the problem lies with the coordinate system used.

One of the main reasons for these problems was that the mathematics needed to work with Einstein's equations and general relativity was not adequately developed. Penrose had done his Ph.D. in mathematics and he developed important mathematical methods to obtain and analyze the results in general relativity. He developed what is known as a Penrose diagram,

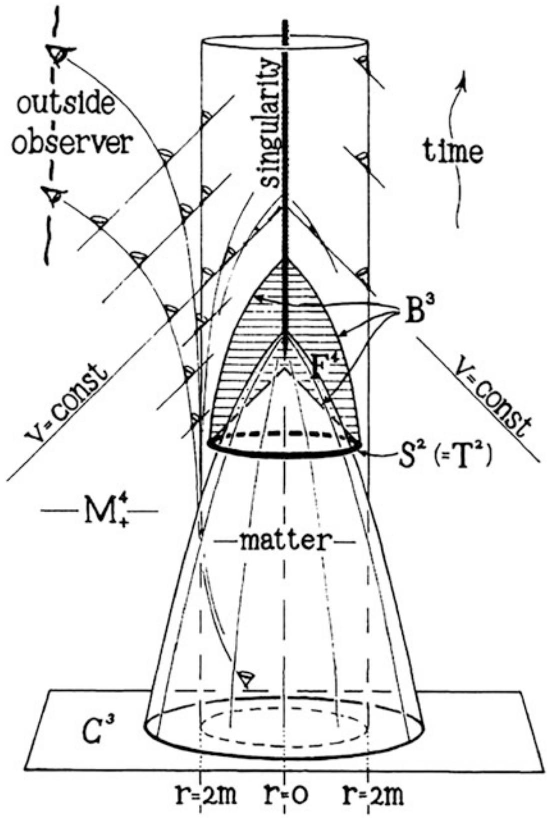


Fig. 6.1 Penrose diagram which shows the inevitability of black hole formation in a stellar collapse

which is a two-dimensional diagram graphically exhibiting the causal relations between different points in space-time. Casual relations describe space-time points which can be connected through cause-effect relations, i.e., a cause occurring at one point can affect what happens later at another point. Kip Thorne of Caltech, a Nobel laureate himself (see Chap. 7), described these contributions of Penrose in the following way: "Roger Penrose revolutionized the mathematical tools that we use to analyze the properties of space-time." In 1965 Penrose published an important work named "Gravitational collapse and space-time singularities" which convincingly proved that a singularity must occur in a gravitationally collapsing star, even when spherical symmetry is absent. Thus, he proved that black holes are a reality and not an artifact of the coordinate system or of the mathematics used in the theory. He also showed that the singularity will be confined well within the event horizon, and outside it, the gravitational field and therefore the curvature of space-time will be large but finite and the region will be visible to the outside world. He summarized

his results in a diagram shown in Fig. 6.1 which depicts the inevitability of the formation and existence of a black hole. Explanation of the diagram is beyond the scope of this book. This 1965 publication of Penrose is considered to be the most important contribution to the general theory of relativity after Einstein's original presentation.

In 1979, Penrose proposed what is known as "strong cosmic censorship", which is still unproved. According to his hypothesis, "a black hole's singularity could not be 'naked'; it had to be confined and forever hidden behind the veil of the event horizon." There would, however, be a visible, strongly curved (though finite) exterior region. He and Hawking did further work to consolidate his earlier work. These results are collectively known as the Penrose–Hawking singularity theorems. These theorems are also of extreme importance for the models of the big bang about which we are going to learn in Chap. 8. Penrose was awarded the 2020 Nobel Prize for his work.

6.3 Nobel Prize 2020: R. Genzel and A. Ghez

The Nobel Prize for the year 2020 was given to R. Genzel and A. M. Ghez "for the discovery of a supermassive compact object at the centre of our galaxy." They shared half the prize amount; the other half was awarded to Roger Penrose.

Credit: Genzel: Nobel Prize Outreach, photo: Bernhard Ludewig

Reinhard Genzel was born in Germany in 1952. After graduating from the University of Frieburg, he obtained his Ph.D. in 1978 from the University of Bonn, after which, he spent a couple of years at the Harvard–Smithsonian Centre for Astrophysics in Massachusetts. In 1981 he joined the University of California, Berkeley as a faculty. In 1986, he became the director of the

Max Planck Institute for Extraterrestrial Physics in Munich. Since 1999 he has also been a part time professor at the University of California Berkeley. He carries out his observational work mainly in the infrared band and uses several instruments on European Southern Observatory telescopes.

Credit: Ghez: Nobel Prize Outreach, photo: Annette Buhl

Andrea M. Ghez was born in 1965 in the USA. After graduating from the Massachusetts Institute of Technology in 1987, she obtained her Ph.D. in 1992 from Caltech. She is a professor of physics and Astronomy in the University of California at Los Angeles and holds the Lauren B. Leitchman and Arthur E. Levine chair in Astrophysics. Ghez is also the Founder and Director of the UCLA Galactic Centre group.

6.3.1 Background

In 1963, Martin Schmidt discovered a radio source, 3C273 (which meant it is object number 273 in the third Cambridge Catalog; see Sect. 4.2) which had a star like appearance. The spectral lines in its optical spectrum were found to be highly redshifted; (see Box 5.1), indicating that the source is moving away from the Earth with a very high speed of about 4.7×10^4 km s^{-1}. As we will see in Chap. 8, the speed of recession of an extragalactic object with respect to the Milky Way gives us the distance of that object from us using Hubble's law, described in that chapter. From its measured redshift, the object 3C273 was estimated to be a couple of billion light years away. An upper limit on its size could be obtained from the time period over which the luminosity of the object varied (see Box 6.1). The size was found to be smaller than about 4 light years, which is 10^4 times smaller than the size of a typical galaxy. From its distance and observed brightness, its luminosity was estimated to be about 1000 times that of our galaxy. Thus, every second, the source was emitting

1000 times as much energy as our Milky Way, and that too from a volume which was a thousand billion times smaller than it.

At first it was thought to be a supermassive star having a mass between 10^5 to 10^9 $M_\odot$. However, it was immediately realized that such stars would be highly unstable and could not exist. Soon, other similar radio sources emitting energy larger than that of an entire galaxy were discovered and all these were named quasars, a short form for quasi stellar radio sources. None of the usual sources of energy could account for the huge amount of energy that seemed to be released from such a small volume of space. Finally, scientists came to the conclusion that the observed amount of energy can only be generated by the conversion of gravitational potential energy of matter falling on a black hole of mass of about 10^8 $M_\odot$. Soon, it was realized that emission from all extragalactic radio sources could be explained on the basis of the presence of supermassive black holes at the centres of galaxies. In the 1970s, it was proposed that the high velocities of stars observed in the central region of elliptical galaxies could be due to the presence of supermassive black holes there.

In 1971, it was speculated that a supermassive black hole could be present at the centre of the Milky Way too. As we have seen in Chap. 4, Jansky had observed radio emissions from the direction of the centre of the Milky Way in 1931. The central region of the Milky Way is in the constellation of Sagittarius. The extended radio source there was later named Sagittarius A. In 1974, a bright and compact radio source which was named Sagittarius A* (Sgr A*) was discovered in the central region of the Milky Way. The name was given as the source was exciting and excited states of atoms are denoted by asterisks. It was suspected that this harbours a massive black hole. One way to confirm the presence of a supermassive black hole in or near Sagittarius A* was to observe the motion of stars and interstellar matter close to it and analyze their trajectories. Such objects should move with very high, relativistic velocities and their motion can be detected via the redshift of their spectral lines or by observing their orbits and their time periods. Applying Kepler's laws to the observed motion, one can estimate the mass of the central object.

Box 6.1 Estimating the Size from Variability

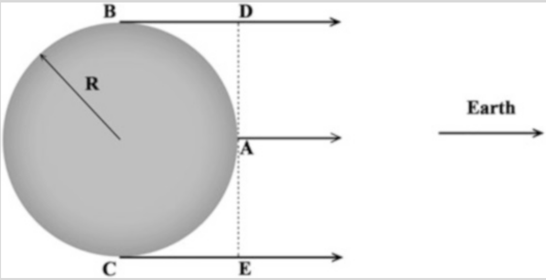
Astronomical sources which are sufficiently close to us, appear like a disc when seen through telescopes. We can easily determine the size of such a source by measuring its angular diameter and the distance of the source from us. However, most astronomical sources are too far away to be seen as discs. They appear to be point sources and we can not determine their sizes directly. For variable sources,

(continued)

Box 6.1 (continued)

i.e., sources whose brightness varies with time, we can determine the sizes by measuring their time period of variation. This can be understood from a simple illustration as follows.

Consider a source which has a spherical surface of radius R shown in the figure. The source is very far from us and appears to be a point source to us. We are interested in estimating its size. If we are situated at a point along the direction indicated in the figure, light from the nearest point on the source, i.e., A, will take least amount of time to reach us, while light from the point B or C will reach us after a time gap close to R/c, which is the light travel time from B and C to D and E respectively, along the lines of sight to us. (Recall that the light rays from A, B and C to us can be assumed to be parallel because of the large distance to the source; and the speed of light in vacuum is constant.) Light from points between A and B and between A and C along the surface of the source, will reach us during this interval. The figure is a two dimensional projection of a sphere, however what we consider below will apply to a spherical source also.



Now let us consider the source to be variable, i.e., its luminosity increases and decreases between certain minimum and maximum values at a time interval of T. In this case, the amount of light emitted by every point on the surface of the source fluctuates simultaneously between the maximum and minimum values with period T. However, the light received by us at a given instant of time from different points on the surface was not emitted at the same time because the light emitted by different points on the surface has to travel different distances and hence takes different amounts of time to reach us as seen above. If T is smaller than R/c, the light reaching us at any given instant emitted by points between A and B, and similarly between A and C, would have left the source when it was at different stages of brightness. Some of these points could have been at their peak brightness while some others could have been at their minimum brightness and the rest would have been at in between states of brightness. The brightness of the source as seen by us, will neither be the minimum nor the maximum but an averaged out value and so will not change with time. Thus, we will not see the source as a variable source. Only if T is greater than or equal to R/c, would the light received by us at a given instant from different points on the surface of the source have been emitted when all the points between B and C along the surface, were in similar, if not the same phase of illumination. For instance, when all these points were emitting higher than average amount of radiation

(continued)

Box 6.1 (continued)
or all of them were emitting less than average radiation. In this case we can see fluctuation in the brightness of the source. Thus, if we see the brightness of a distant source varying appreciably over a time T, we can be sure that its size is smaller than or equal to $T \times c$. This is one method of estimating the sizes of distant variable sources.

6.3.2 Award Winning Work: Discovery of a Supermassive Compact Object at the Centre of the Milky Way

The centre of our galaxy is at a distance of about 26000 light years from us. There are two major problems for observing objects present near it. Firstly, the lines of sight to these objects pass through the disc of the Galaxy which contains most of its matter, and so the photons emitted by these objects have to travel through a tremendous amount of gas and accompanying dust. Because of this, when we are observing at optical wavelengths, there is a very high probability of the photons getting absorbed before reaching us, and only one in 10 billion photons coming from these objects will be able to reach us. It is therefore impossible to observe anything close to the Galactic centre using optical telescopes. The situation improves dramatically if one chooses to observe in the infrared, as the probability of scattering and absorption of infrared photons by interstellar gas and dust is small. At those wavelengths, nine out of ten photons are able to reach us. For this reason, the two groups led by Genzel and by Ghez chose to observe at 2.2 microns, a micron being a millionth of a metre. Genzel's group used the VLT. This belongs to the European Southern Observatory and consists of four telescopes, each having mirrors of 8.2 m diameter. Ghez's group used the twin KECK telescopes in Hawaii, each telescope having a mirror of 10 m diameter. Both groups have now been working on this project for more than 25 years.

The second problem while observing objects near the Galactic centre is that of angular resolution. The Galactic centre is at a large distance from us, and the entire region containing a central supermassive black hole and nearby stars that one would like to study, subtends a very small angle at our telescopes. Very high resolution telescopes are needed to see individual stars in the region, and to determine their motions accurately. For ground based telescopes, the resolution is poor due to scintillation caused by the Earth's atmosphere (see Sect. 4.3.2).

Both groups used innovative techniques to overcome this effect. From the nineties to about 2004, they both used the techniques of speckle imaging and speckle interferometry. In the speckle imaging method, a large number of very short exposure (of about one tenth of a second) images are taken. The atmospheric conditions can be assumed to be constant during any one exposure. The images can then be combined using special technique, called "shift and add", which essentially brings the images of the source in different exposures to a single point which is common to all images, and then adds them. This generates a high resolution bright image of the source as the effect of atmosphere gets averaged out. Speckle interferometry uses a mask with holes to be placed in front of the telescope. This produces multiple images and allows interferometric observations, which again improves the resolution of the final image.

After 2004, the technique of adaptive optics was used by both the groups. This technique measures the atmospheric fluctuations near the source position in real time by observing its effect either on a nearby bright star with an accurately known position, or on an artificial star produced at a nearby fixed position on the sky with the help of a laser beam projected in the atmosphere from the surface of the Earth. The measured effect of the atmospheric disturbance is then compensated in the observations of the source by instantly changing the shape of the telescope mirror which is deformable. This takes place in real time with the help of computers, rendering the observed image of the source free from any scintillation caused by the atmosphere. Using this technique, high resolution images of the Galactic centre could be obtained and the motion of stars very close to the Galactic centre could be monitored.

Both teams were able to resolve several stars, which were found to be orbiting the central mass, and were able to measure their motions. The time periods of orbital motion of some of the stars were small. For one star named as S2, the period was as small as 16 years and so even its complete orbit could be observed. The image of the central region of the Milky Way and the observed positions of the star S2 at different times are shown in the left and right panels of Fig. 6.2. The continuous line in the right panel shows the best fit to the observed positions of S2. The orbit is an ellipse with Sgr A* near one of its foci. At closest approach to Sgr A*, S2 was about 14 billion km away from it and it was moving at a speed which is about 3% of the velocity of light. The orbits of several other stars close to Sgr A* were also observed. These orbits were analyzed by applying Kepler's laws and it was discovered that Sgr A*, around which the stars were orbiting, had a mass of $4.31 \times 10^6\ M_\odot$. The mass measurement is believed to be accurate to 1%. This mass was present in a spherical volume having a diameter roughly equal to 17 light hours.

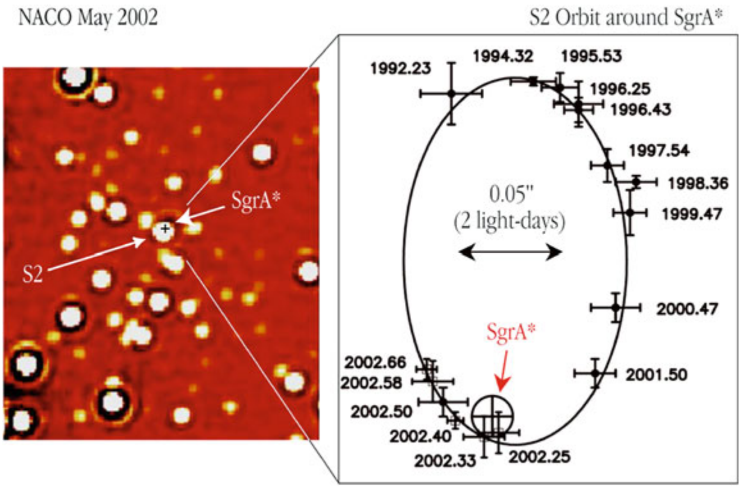


Fig. 6.2 The left panel shows the central portion of the Milky Way. The position of Sgr A* and that of the star S2 are indicated. The right panel shows observed positions of S2 over several years and the best fit ellipse to the observed positions. **Credit:** ESO

This is roughly 100 times the Schwarzschild radius of a black hole having the measured mass. The upper limit on the size of Sgr A* could later be improved on the basis of its variability (see Box 6.1) and an upper limit of about five times the Schwarzschild radius could be obtained. This is ample proof that Sgr A* is indeed a supermassive black hole.

In addition to proving the presence of a supermassive black hole at the centre of our galaxy, the observations of the two groups could also test Einstein's general theory of relativity in two ways. First was the observation that the elliptical orbit of S2 as indicated by the observations was not a closed orbit, i.e., the position of the star at a given time and its position after it completes one revolution around Sgr A* were not identical. This indicated that the orbit of S2 is precessing. As seen in Sect. 6.2.1, such precession of the planet Mercury has been observed and is consistent with the predictions of Einstein's theory. Such precession has also been observed in the case of binary pulsar systems. The precession of the orbit of S2 was the first such precession to have been observed in relation to a supermassive black hole. The other test was that of gravitational redshift (see Sect. 6.2.1). The light emitted by S2 has to overcome the gravitational pull of the supermassive black hole in Sgr A*. During its motion along its orbit, its distance from Sgr A* changes and so does the strength of the gravitational field to be overcome by the photons emitted by S2. Thus, the radiation coming from S2 should show changing

redshift. This effect has indeed been observed. The results have thus provided proof of the general theory of relativity under very strong gravitational field of a supermassive black hole. It may be noted that in this measurement, the redshift of light due to the large velocity of S2 due to the Doppler effect was properly accounted for. Genzel and Ghez were awarded the Nobel Prize for the year 2020.

7

Gravitational Waves

7.1 Introduction

The Nobel Prizes in the category of gravitational waves were given in the years 1993 and 2017 to five astronomers: Russell Alan Hulse, Joseph Hooton Taylor Jr., Barry Clark Barish, Kip Stefan Thorne, and Rainer Weiss. Hulse and Taylor received the prize for the year 1993 for the discovery of binary pulsar, their observations providing indirect proof of the presence of gravitational waves. Their work started in 1974 and they announced their proof for the existence of gravitational waves in 1978. Barish, Thorne, and Weiss received the prize for the year 2017 for their contributions in setting up the gravitational wave observatory and for the first actual detection of gravitational waves in 2015 which was announced in 2016. The 2017 prize was perhaps the fastest a work got recognition in the form of a Nobel Prize, as even though the work to detect gravitational waves started decades ago, the actual detection was announced only an year before the year of the award. This underlines the immense importance of the discovery and is a tribute to the enormous amount of planning and work that preceded the actual detection.

P. Khare, *Nobel Prizes in Astronomy*, Astronomers' Universe,
https://doi.org/10.1007/978-3-031-29639-0_7

7.2 Nobel Prize 1993: R. A. Hulse and J. H. Taylor Jr.

The Nobel Prize for the year 1993 was awarded to Russell A. Hulse and Joseph H. Taylor Jr. "for the discovery of a new type of pulsar, a discovery that has opened up new possibilities for the study of gravitation."

Credit: Hulse: By ENERGY.GOV—HD.3A.054, Public Domain, https://commons.wikimedia.org/w/index.php?curid=35935806

Russel Allen Hulse was born in 1950 in the USA. He graduated in 1970 from the Bronx High school of Science, New York and went on to obtain his Ph.D. in 1975 from the University of Massachusetts, Amherst under the guidance of Taylor. From 1977 he worked in the plasma physics Laboratory of Princeton University for several years. In 2003 he joined the University of Texas as a visiting professor and later became a full time professor. He went on to become the Founding Director of the UT Dallas Science and Engineering Education Center.

Joseph Hutton Taylor Jr. was born in the USA in 1941. After graduating in physics in 1963 from Haverford College, Pennsylvania he obtained his Ph.D. in astronomy in 1968 from Harvard University. He worked as a faculty at the University of Massachusetts, Amherst from 1969 to 1981, and became the Associate Director of the Five College Radio Astronomy Observatory there. He later joined as a professor of physics at Princeton University and during his 25 year career there, he was the James S. McDonnell professor of physics and also served as the Dean of the physics department for a period. He became professor emeritus there in 2006. He was the first person to discover a pulsar outside the Cambridge group (UK) which had discovered the first pulsar.

7.2.1 Background

We have seen (see Box 2.1) that in any wave, the value of some entity increases and decreases periodically. In the water waves in a pond, it is the height of water particles, from their mean position of rest, which changes with time and with distance as the wave propagates, while in electromagnetic waves, the varying entities are the electric and magnetic fields. In the case of gravitational waves it is the gravitational field that changes with time and distance. The existence of such waves was first predicted by Einstein in 1916, using his new theory of gravitation, general relativity (see Sect. 6.2.1), which he had announced a year earlier. In this theory, the force of gravity is explained as being due to the curvature of space-time. The presence of matter and energy generates curvature in space-time. This affects the way particles and light rays move. The greater the curvature, the greater is the gravitational force. Gravitational waves are periodic changes in the gravitational field, i.e., in the curvature of space-time. This means they are undulations or ripples in the fabric of space-time, just like water waves are ripples on the surface of water.

However, for gravitational waves, the situation is more complicated than that for water waves or electromagnetic waves. This is because, in Einstein's theory, the gravitational field itself is at the basis of the definition of space-time. It is therefore difficult to distinguish between real changes in gravitational field from changes merely in the way space and time are measured. The confusion was such that Einstein himself did not fully believe his own result, and in fact declared in 1936 that the waves did not exist. He changed his ideas after peers pointed out the fallacy in his arguments. Even though a number of scientists contributed to the final acceptance of gravitational waves as a natural outcome of Einstein's theory of general relativity, Einstein is credited with their prediction.

Like electromagnetic waves, gravitational waves can travel in vacuum and they travel with the same speed as that of light. These waves also get redshifted or blueshifted as per the motion of the source of the waves. One major difference between electromagnetic waves and gravitational waves is that electromagnetic waves get easily absorbed by matter and so can not travel long distances through matter, while gravitational waves can travel long distances. This is because gravitational interaction is 10^{36} times weaker than electromagnetic interaction (see Sect. 2.2.1) and the probability of gravitational waves getting absorbed by matter is extremely small. Thus, gravitational waves emitted by very distant sources are able to reach us.

How are these waves emitted? Accelerated motion of a mass emits gravitational waves. For example when we move our hand, or walk, these waves are generated. The objects that we come across in our day to day life all have very small masses and have rather small accelerations, hence the gravitational fields of these objects, and the fields in the gravitational waves generated by them, are very weak. Even the gravitational waves generated by planets and stars like the Sun are very weak. By "weak", we mean that the magnitude of the gravitational field in these waves is extremely small. As we will see below, the gravitational wave detectors typically measure the periodic displacement of so called test masses which are masses free from any external forces, under the influence of the periodic gravitational field of a wave. The gravitational fields in these waves and the gravitational interaction both being weak, the displacement caused by them in the test masses is extremely small and it is difficult to detect these waves.

It is beyond our present capability to construct a machine on the Earth which will be able to generate gravitational waves which can be detected using existing technology, and we have to turn to much larger masses and much higher accelerations than found on the Earth or even in the solar system. For gravitational waves to be detectable, the gravitational field that they carry to the Earth should be sufficiently strong to be able make changes (in a detector) which can be measured and interpreted as gravitational waves. For this, the waves have to be produced by objects which are massive and highly compact, so that their gravitational field is very strong. Moreover, the objects have to undergo large acceleration, so that they can radiate strong gravitational waves. Such objects are rare and they are more likely to be found in distant regions of our galaxy and in other galaxies. Gravitational waves emitted by these sources have to travel over great distances to reach the Earth, so their energy is spread over a very large sphere, centred at the source, whose radius is equal to the distance between the source and the Earth. Therefore, however strong the gravitational waves at the source may be, we receive only a very weak flux at the Earth, which makes the detection very difficult.

Most astronomical sources have large mass, but not all of them are compact and/or undergo significant acceleration. The highest accelerations are obtained by either the material of exploding stars or by the motion of compact stellar remnants in binary systems. We can divide the possible important sources of gravitational waves in 4 categories. (1) Burst sources, like supernova explosions and Gamma-ray bursts (see Box 7.1), which are short-duration one off events; (2) Continuous sources like spinning, compact objects e.g., pulsars, which can emit gravitational waves at the same frequency as its frequency of rotation, for a long duration; (3) Binary sources with the two components being

compact objects, i.e., white dwarfs, neutron stars, or black holes; (4) Stochastic background sources which are collections of a large number of weak sources, which cannot individually be detected but can be collectively observed as a background of gravitational waves.

Box 7.1 Gamma-Ray Burst

Gamma-ray bursts are extremely high energy explosive events. In such a burst, first there is an explosive release of Gamma-Rays, which is followed by electromagnetic radiation of various kinds, including X-rays, optical radiation, and radio waves. They were first detected in 1967 by satellites flown for the purpose of detecting man made nuclear explosions. The data was declassified only in 1973. By now thousands of Gamma-Ray bursts lasting for various durations, ranging from a fraction of a second to hundreds of seconds and longer, have been observed. These are of extragalactic origin. The amount of energy emitted in such a short time is equivalent to the total energy emitted by a star like the Sun in about a trillion years.
The bursts are mainly of two types: (1) Long duration bursts which last for longer than about 2 s, whose origin is believed to be in core-collapse supernovae of rapidly rotating stars. Such an event will emit gravitational waves of the type described above. (2) Short duration bursts of less than about 2 s, which are believed to originate in the merger of two neutron stars in a binary system. Gravitational wave detection from a merger of this type, followed immediately by the detection of the associated Gamma-ray burst has already been observed, as we will see below.

7.2.2 Award Winning Work: Discovery of a Binary Pulsar

In 1974 Taylor and Hulse, who was then Taylor's graduate student, carried out a very sensitive survey for observing new pulsars. It was then only seven years since the discovery of the first radio pulsar, and only about 100 pulsars had been observed. Just to compare, more than 2700 radio pulsars have been observed until the end of 2021. Hulse and Taylor used the largest radio telescope available at that time, which was located at Arecibo in Puerto Rico. They also used computers to aid their search, and their detection capability was ten times higher than that of earlier observations. Until 1974, it had not been possible to determine the mass of any of the then known pulsars. We have seen in Box 3.2 that it is possible to determine the masses of stars which are members of certain types of binaries. Hulse and Taylor were hoping to find a pulsar which is a member of a binary, so that they would be able to determine its mass. They discovered a pulsar on July 2, 1974. It had a rotation period of

59 ms and was the second fastest rotating pulsar at that time, the fastest being the Crab nebula pulsar with a period of 39 ms. It was named PSR B1913+16 according to its position, i.e., its coordinates in the sky.

As we know, every time the pulsar beam is directed towards the Earth, we receive a pulse. Thus, the pulse period is same as the period of rotation of the neutron star. A neutron star, being massive and compact, is very robust and its rotation period remains constant for a long time. There is a very slow increase in this period due to the fact that the pulsar loses energy due to emission of radiation. This was indeed true for all the 100 pulsars observed till then. Surprisingly, Taylor and Hulse found out that the period of the PSR B1913+16 was changing from day to day. This could not have been due to a change in the rotation rate of the neutron star. After a careful study of the observed data they concluded that the apparent change in rotation period was occurring due to the fact that the pulsar was a member of a binary system with the two stars going round each other every eight hours.

Hulse and Taylor continued to study B1913+16 for many years with the Arecibo radio telescope. Over the years, accurate information about the masses of the two stars and their orbits has been obtained. Even though the masses of any two stars in a binary system can be obtained in principle, in practice, there are difficulties like determining the inclination of the plane of the orbit to the line of sight. These difficulties do not allow mass determination for all binaries. In this case of binary neutron stars, however, the masses could be determined with unprecedented accuracy because the other general relativistic effects that are significant due to the very high gravitational fields, e.g., gravitational time dilation (see Sect. 6.2.1) experienced by the pulsar, could be observationally determined. The masses of the two neutron stars in the binary system are $1.4398\,M_\odot$ and $1.3886\,M_\odot$. The orbit of the pulsar has been observed to precess due to general relativistic effects (see Sect. 6.2.1). The rate of precession is much higher than that of the orbit of Mercury, due to the high gravitational field experienced by the pulsar. It is 4.226598 deg per year while that for Mercury is 575 arc sec per century. The binary has a very eccentric orbit, so each star moves around the companion star along a path which is a narrow ellipse, as shown schematically in Fig. 7.1. The minimum distance between the two stars, which is known as the perihelion distance, is about 746,600 km which is comparable to the size of the Sun–which is 700,000 km–and is much smaller than the maximum distance between them which is about 3.15 million km. The minimum distance is only about 1/50th of the distance between the Sun and the Earth. If the companion was an ordinary star, it would have collided with the pulsar when the two were closest to each other and the binary would have been disrupted. Also, if it were an extended object like an

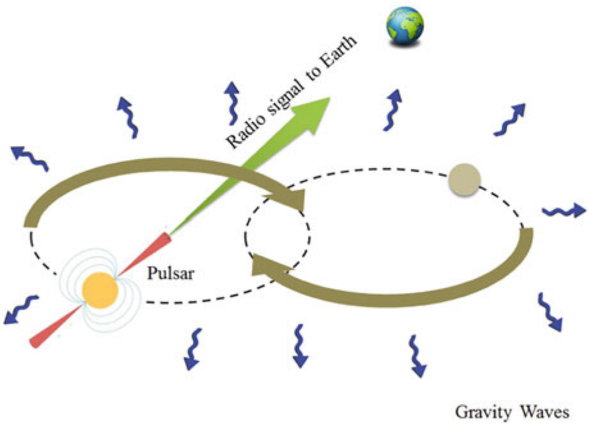


Fig. 7.1 Artistic impression of the binary pulsar discovered by Hulse and Taylor. One of the two stars is a pulsar. The other too is a neutron star but not a pulsar. The pulsar emits radio waves and other radiation in narrow cones, while the binary emits gravitational waves in a broad pattern

ordinary star, then as the pulsar went round it, the pulsar would be eclipsed, but this was not observed. Such arguments led to the understanding that the companion star is also a very compact star and must be a neutron star. We do not see it as a pulsar because either the radio beam from the companion does not sweep past the Earth, or it is not a pulsar. The radii of the two stars have been estimated to be about 10 km.

The fact that both components of the binary are so compact is important in the analysis of the system. Due to the small distance between the two components, i.e., high gravitational field experienced by both, and their high velocities, Newton's theory of gravity does not hold good for the system, and general relativistic effects come into play as noted above. Such high gravitational fields and high velocities are rarely seen among stars and therefore, the binary pulsar has proved to be an ideal laboratory for the study of gravitational effects.

Soon after the discovery of the binary pulsar, it was pointed out by Robert Wagoner that the observation of change in the binary rotation period with time should allow us to test the presence of gravitational waves. The binary pulsar is very compact and the gravitational force between the two neutron stars is very strong. The orbit of the pulsar around the companion is highly elliptical in nature and the accelerations that the neutron stars undergo are also large. As seen in Sect. 7.2.1, such a binary emits significant amounts of gravitational waves due to the high accelerations of its components. The emission results in loss of energy by the binary, and consequently the binary shrinks in size. The two neutron stars come closer together and because of the principle of conservation of angular momentum, revolve faster around each other, leading to a decrease in the period of the binary. A measurement of the change in

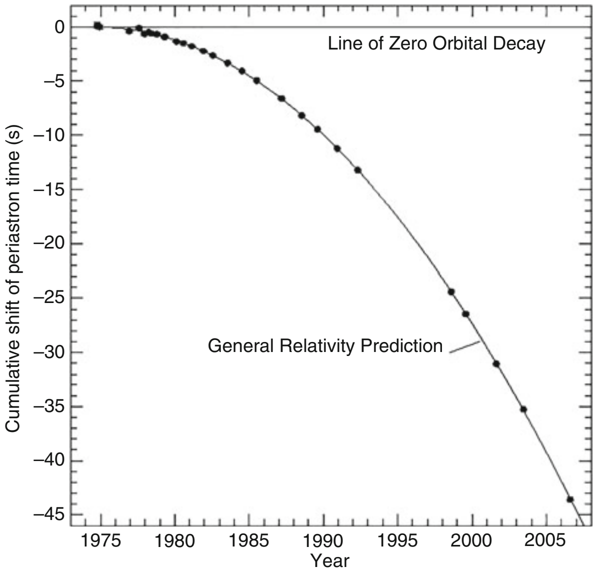


Fig. 7.2 The change in rotation period of the binary B1913+16 over a span of 30 years from 1975 to 2006, as predicted by Einstein's General theory of relativity, is shown by the continuous curve, and the observed values are shown by dots. The *x*-axis shows the time of observation and the *y*-axis shows the decrease in time taken for successive revolutions, which indicates shrinking of the orbit. If gravitational waves were not being emitted, the graph would have been a straight line, as shown at the top. **Credit:** J. M. Weisberg, D. J. Nice and J. H. Taylor, Asrophys. Journ. **722**, 1030, 2010

their period of revolution would allow the measurement of the rate at which they are losing energy through the emission of gravitational waves. This could be compared with the value predicted by Einstein's theory. If the expected and observed values happen to be equal, it would provide a validation of the existence and emission of gravitational waves.

Hulse and Taylor continued their work on the observations of the binary pulsar system and in 1978, four years after its discovery, announced that they had indeed found a decrease of the binary period which was consistent with the prediction of Einstein's theory. The theoretical general relativistic value was obtained using the measured parameters of the binary pulsar. The cumulative shift over many years of observation reported in 2010, is shown in Fig. 7.2. The continuous curve shows the cumulative shift with time as predicted by Einstein's theory, while the dots are the observed values. The dots fall very closely on the predicted line, with any observed departure from the line being too small (smaller than 1%) to be seen. The emission of gravitational waves is therefore very well validated. This was the first time that the gravitational waves

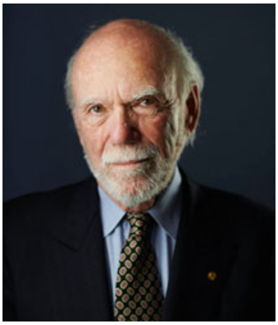
predicted in 1916 were seen to really exist. The binary pulsar also provided an ideal laboratory for studying general relativistic effects, due to the high gravitational field between the two stars. Hulse and Taylor were awarded the Nobel Prize for the year 1993 for their work.

7.3 Nobel Prize 2017: R. Weiss, B. C. Barish, and K. S. Thorne

The Nobel Prize for the year 2017 was given to B. C. Barish, K. S. Thorne and R. Weiss "for decisive contributions to the Laser Interferometer Gravitational-Wave Observatory detector and the observation of gravitational waves." Weiss received half the prize amount; the other half was shared by the Barish and Thorne.

Credit: Nobel Media, photo: A. Mahmoud

Rainer Weiss was born in 1932 in Germany. His family migrated to the US in 1939. He graduated from the Massachusetts Institute of Technology (MIT) in 1955 and continued there to get his Ph.D. in 1962. After working as a faculty for two years each at the Tufts and Princeton Universities, he joined as a faculty in MIT in 1964 and retired from there in 2001. He has been an adjunct professor at Louisiana State University since. He was a co-founder and an intellectual leader of the Cosmic Microwave Background Explorer (COBE) mission.



Barry C. Barish was born in 1936 in the US. He obtained his Bachelor's degree and Ph.D. in 1957 and 1962 respectively from the University of California, Berkeley. He joined Caltech in 1963 as a research fellow and subsequently became professor there. In 1991 he became the Linde professor of physics and after 2005, he has been the Linde professor of physics, emeritus there.

Kip Thorne was born in 1940 in the US. He obtained his Bachelor's degree from Caltech in 1962 and Ph.D. from Princeton University in 1965. He joined Caltech as a faculty in 1967 and continued there till 2009. He became the William R. Kenan, Jr. Professor in 1981, and the Feynman Professor of Theoretical physics in 1992, after which he became Feynman professor emeritus. He resigned in 2009 to devote his time to other pursuits like writing and movie making. He was an executive producer and scientific adviser for the science fiction movie "Interstellar" by Christopher Nolan. He is a coauthor with John Wheeler and Charles Misner of the book "Gravitation", published in 1973, which has been very influential.

7.3.1 Background

Till about 400 years back, cosmic objects could only be observed through the visible radiation emitted by them, and that too only with naked eyes. Galileo's pointing his telescope at the night sky changed this. Improved detectors and bigger telescopes gave us a huge new insight into the Universe. Slowly we built detectors and telescopes for other bands of electromagnetic radiation, some of which are described in Chap. 4. However, till the year 2016, electromagnetic radiation remained the only major source of information about the cosmic objects. The situation was as if there was only one "window" available to us for observing the Universe, that of electromagnetic radiation. Gravitational waves were considered to be another, different type of source for obtaining information about the Universe. These waves are not electromagnetic waves and so can provide an entirely different type of information as if offering us a peek at the Universe through another "window". Through gravitational waves, we can observe objects/events in the Universe from which electromagnetic waves are not emitted, such as events involving black holes. We have seen that these waves do not easily get absorbed by matter, and so can travel over much larger distances compared to electromagnetic radiation. These would, in principle, allow us to observe more distant and different kinds of objects in the Universe than the objects that we are able to observe through electromagnetic radiation. Thus, direct observation of gravitational waves promised completely new information about the so far unobserved aspects of the Universe. However, as we have seen in Sect. 7.2.1, these waves are much more difficult to detect because of their weak gravitational interaction. Their detection is extremely challenging and a huge amount of resources, hard work and time were needed to accomplish their successful detection.

A gravitational wave has gravitational fields which oscillate as a function of space and time. Particles lying along the path of such a wave experience these gravitational fields, and as a result, get displaced from their positions of rest. Their displacements would be periodic because of the periodic nature of the gravitational fields in the wave. By observing such oscillations of test particles (test masses), we can detect a gravitational wave. The displacements of test masses on the Earth caused by gravitational waves of cosmic origin are expected to be extremely small, about 8500 times smaller than the radius of a proton. It is very difficult even to conceive the measurement of such small displacements. The first gravitational wave detectors based on this idea were built by Joseph Weber of the University of Maryland in the early 1960s. He constructed aluminum bars of different lengths and diameters. He

hoped that there would likely be a resonance between the frequency of the gravitational waves expected from compact cosmic sources and the natural frequency of oscillation of particles of some of the bars. Resonance enhances the displacements of particles, making it easier to detect them. If a gravitational wave passed through a bar, the bar would start vibrating. The vibration, though, would be extremely feeble and hard to detect. For this purpose, Weber placed piezoelectric crystals on the surfaces of the bars. These crystals have the property that they produce electric voltages when compressed or stretched. The vibration of the bar would generate tiny voltages in the crystals which could be measured using sensitive electronic instruments. However, these detectors and similar detectors constructed by other scientists failed to detect gravitational waves.

The next gravitational wave detectors used interferometers. We have learnt about interference in Sect. 4.2.1 (see Box 4.1). This phenomenon was used by Michelson in an optical instrument that he developed in 1881 and is named after him as Michelson's interferometer. A very simple version of the Michelson interferometer is shown in Fig. 7.3. A beam of light emitted by a source at the left of the diagram travels to a mirror which is known as a beam splitter. This mirror allows half the light of the beam to pass through it, towards mirror M_1, while the other half of the light is reflected towards mirror M_2. The light striking the mirrors M_1 and M_2 is reflected back towards the beam splitter, and the combined beam then reaches a light detector at the bottom of the figure. The two paths of the beam between the beam splitter and M_1, and the beam splitter and M_2, are said to be two arms of the interferometer. When the two beams combine together, the outcome depends on how the corresponding

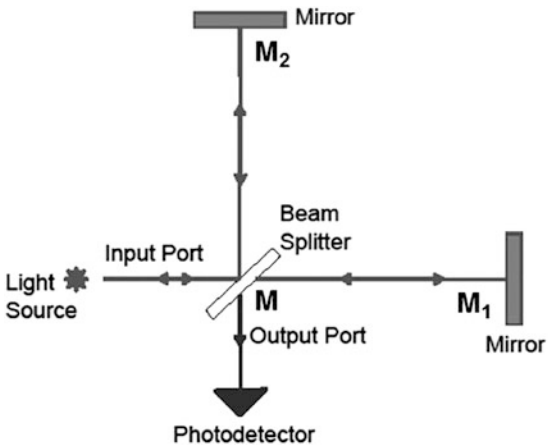


Fig. 7.3 A simple depiction of Michelson's interferometer described in the text

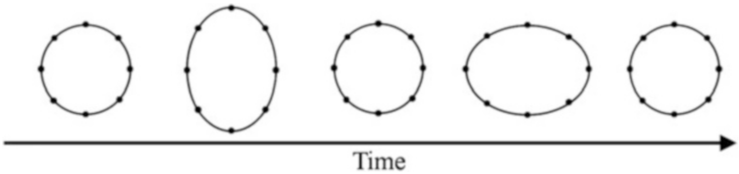


Fig. 7.4 The effect of a gravitational wave on a set of free point masses arranged in a circle. The gravitational wave travels into the plane of the paper, in a direction perpendicular to the plane

waves are matched. When the distances traveled by the two beams are exactly the same or differ by an integral multiple of the wavelength, the waves combine together constructively (see Figure A in Box 4.1), producing a bright spot at the detector. If the distances traveled differ by odd multiple of half the wavelength, then the two waves combine destructively, producing a dark spot. Depending on the exact configuration of the mirrors M_1 and M_2, the combined beam can produce interference pattern consisting of concentric light and dark circles, or of light and dark straight lines at the detector. If either M_1 or M_2 is moved slightly, the distance traveled by one of the beams changes, leading to a change in the observed interference pattern. Michelson's interferometer is essentially used to accurately measure lengths and changes in lengths.

The concept of the interferometer provides a way to detect gravitational waves. To see this, let us consider tiny particles which are arranged in a circle in space, as shown on the left in Fig. 7.4. Now suppose a gravitational wave passes in a direction perpendicular to the circle into the plane of the paper. The change in the geometry produced by the wave is such that there is alternately expansion and contraction of distances along the horizontal and vertical directions in the plane of the paper: when there is contraction along the horizontal direction as shown by the next shape to the right of the circle in the figure, there is expansion along the vertical direction. As time passes, and the gravitational wave propagates forward, the directions of contraction and expansion are exchanged as shown by successive shapes towards the right. For particles which are not along the horizontal or vertical directions, expansion or contraction takes place depending on the direction in which they are located. The net effect is that the circle of particles changes shape periodically: it alternately becomes an ellipse first stretched along the y-direction and then stretched along the x-direction, as shown in the figure.

Now imagine that there are only two particles on the circle and mirrors M_1 and M_2 of the Michelson's interferometer are attached to them, with the beam splitter M placed at the centre, as shown on the left in Fig. 7.5. The

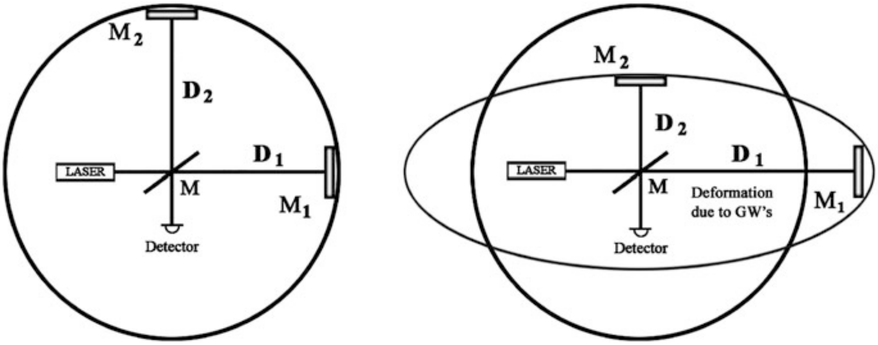


Fig. 7.5 On the left are shown the mirrors of a Michelson's interferometer in their normal state, when they are on the circle. The mirrors M_1 and M_2 are at equal distance from the mirror M. On the right the configuration when a gravitational wave is passing through is shown. M_1 has moved away from the centre of the circle, while M_2 has moved towards the centre. The mirrors are now on an ellipse. The change in distances of the two mirrors from the centre causes a change in the interference pattern. **Credit:** Kaushal Sharma and Ajith Parameswaran

distances traveled by the two beams, namely 2 D_1 and 2 D_2 are equal and the interference is constructive so a bright spot or a bright fringe is seen at the centre of the detector. Now suppose a gravitational wave passes through the circle of particles, so that it is stretched to form an ellipse as shown on the right of the figure. The mirror M_1 is now farther from M than it was in the absence of the wave, while M_2 has moved closer to M, so D_2 is smaller than D_1. The distances traveled by the two beams are therefore different and for a certain difference in the path length, the interference is destructive and a dark spot or a dark fringe is seen at the centre. As the mirrors move back and forth the path lengths change periodically, which can be detected by the periodically changing interference pattern. The changing interference pattern becomes a signature of the passage of a gravitational wave.

7.3.2 Award Winning Work: LIGO and Detection of Gravitational Waves

The use of interferometers for detecting gravitational waves was first proposed in 1962 by the Russian physicists Michael Gertsenshtein and Vladislav Pustovoit, and later independently by Joseph Weber and Rainer Weiss in the USA. Weiss fully developed the concept. The interferometer was to use laser light instead of ordinary light that is used in Michelson's interferometer, and the

observatory was called Laser Interferometric Gravitational wave Observatory, or LIGO in short. It was decided to have large interferometers at two sites, one at Hanford, Washington State and the other at Livingston in Louisiana, both in the USA, about 3000 km apart. Weiss also made major contributions to the development of LIGO. The construction of the LIGO detectors was jointly taken up by MIT and Caltech.

Barish became the Principal Investigator of the LIGO project in 1994 and in 1997 became its Director. He led the efforts in the final design, approval for funds, and the construction and commissioning of the two LIGO interferometers. He also set up the LIGO Science Collaboration in 1997, whose mission is to detect gravitational waves, use gravitational waves to explore fundamental physics of gravity, and to develop gravitational wave observations as a tool for astronomical discovery. Weiss was its first spokesman. The collaboration now has over 1200 members from over 100 institutions in 18 countries.

Thorne developed the formalism for the analysis of the generation of gravitational waves. He provided theoretical support for LIGO which included determining the type of sources LIGO should target and several other technical aspects of LIGO. He developed the mathematics that would be needed to analyse the observations. His contribution to the LIGO project is enormous. He is a co-founder of the LIGO project.

It was mainly the efforts of these three, undoubtedly aided by hundreds of other scientists, which culminated in the setting up and working of the two LIGO detectors that resulted in the first ever detection of gravitational waves on 15th September 2015. We will learn more about this below.

LIGO is essentially a Michelson interferometer for detecting gravitational waves as described above. A typical Michelson interferometer used for optical measurements has an accuracy of up to 10^{-12} m. The accuracy required for measuring the displacements caused by gravitational waves is about 10^{-19} m. This is achieved by adapting a series of technical innovations designed to enhance the ability to measure the displacement of the mirrors and to eliminate or minimize spurious vibrations, which we can term as noise. Some of these measures are: (1) increasing the length of the arms of the interferometer to 4 km; (2) using a powerful laser as the source of the light beams; (3) using extra mirrors to increase the intensity of light and to increase the distance traversed by the beam through repeated reflections; (4) placing the entire interferometer in a high vacuum enclosure to reduce noise and (5) using very sophisticated suspensions for holding the mirrors to overcome seismic noise. An aerial picture of the LIGO detector in Hanford is shown in Fig. 7.6.

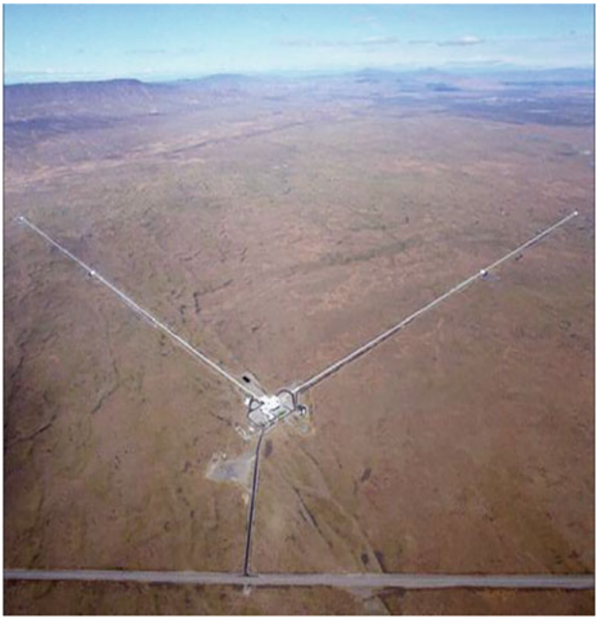


Fig. 7.6 An aerial view of the LIGO facility in Hanford, Washington State, USA. The two, 4 km long arms of the detector can be seen. **Credit:** CALTECH/MIT/LIGO Laboratory, https://www.ligo.org

The first LIGO detectors, now called the initial LIGO or iLIGO, started observing in 2002 and continued observations for a decade without any success. The detectors were upgraded to Advanced LIGO or aLIGO which were ready in 2015. In September 2015, a signal was received by both the aLIGO detectors with a time gap of 7 ms. The data were analyzed thoroughly to rule out the possibility of any spurious signal, and it was concluded that the signal was indeed due to a gravitational wave. The upper panel of the Fig. 7.7 shows the signal detected at Hanford. In the lower panel of the figure, the signal detected at Livingston is shown, together with the signal detected at Hanford, so that the two signals can be compared. The time from the start of the signal is plotted along the horizontal axis, while the vertical axis shows the strain, that is the change in length of the detector arm divided by its steady length in the absence of the wave. The time covered by the entire signal is about 0.2 seconds. The time interval between successive peaks in the signal becomes shorter, i.e., the frequency of the wave increases. So does the height of the peaks, i.e., the amplitude of the wave increases with time. After reaching a maximum, the amplitude decreases rapidly over successive cycles. This is known as the ringdown.

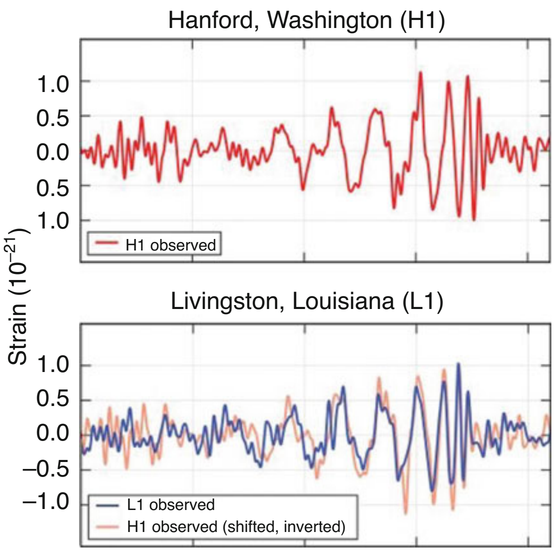


Fig. 7.7 The upper panel shows the signal from the first detected gravitational wave as detected at Hanford. The lower panel shows the signal detected at Livingston, along with the signal detected at Hanford for comparison. **Credit:** B. P. Abbott et al. PRL **116**, 061102, 2016

From the nature of the signal with its increasing frequency and amplitude; it follows that the source is very likely to be a binary system, with the two components rapidly going round each other. Such a binary emits gravitational waves and thus loses energy continuously, which causes the two components to spiral towards each other. This results in a decrease in the distance between them and an increase in the frequency with which they go around each other. The frequency of the gravitational waves emitted by the binary system is primarily twice the frequency of revolution of the binary. After reaching the highest peak, the two objects merge into each other and the resulting object slows down (ringdown) due to continued emission of gravitational waves.

A detailed analysis of the data showed that the two components of the binary were black holes with masses of 36 $M_{\odot}$ and 29 $M_{\odot}$. After merging they became a single black hole with a mass of 62 $M_{\odot}$. Before merging, the total mass was 65 $M_{\odot}$. The difference of 3 $M_{\odot}$ was emitted in the form of gravitational waves. This binary was at a distance of 1.34 billion light years from us, and it was the first binary black hole system to be discovered. These are also the first observed black holes with the masses in the range observed. Before this, the only black holes to be observed were either in X-ray binaries (see Sect. 4.4.2) and had masses less than 10 $M_{\odot}$, or supermassive black hole in the centres of galaxies (see Sect. 6.3.2). Such a black hole binary merger could not have been detected earlier, as black holes do not emit any electromagnetic radiation. The discovery was announced on 11th February, 2016. Such was the importance of the discovery that the Nobel Prize for 2017 was awarded for it.

The aLIGO detectors have been operational since their first detection and have completed 3 observing runs. They have observed a total of 90 sources out of which 2 are binary neutron star mergers, 3 are neutron star-black hole binary mergers, and the rest are black hole binary mergers. The black hole masses range from about 5 to 100 $M_{\odot}$. The binary neutron star mergers are interesting as they also emit electromagnetic radiation while no electromagnetic radiation is emitted by the black hole binary mergers. The gravitational waves from the first neutron star merger event were observed in 2017. The electromagnetic radiation from this event has also been observed in several bands and has provided important information about the nature of the merger. Several elements heavier than iron have been observed in these sources, confirming the neutron star mergers as yet another site for nucleosynthesis of heavy elements. Gravitational wave detections have provided unprecedented data about black holes and neutron stars, and will help astronomers understand them much better.

Currently, four additional gravitational wave detectors are functional. These are VIRGO in Italy, GEO600 in Germany, and TAMA300 and KAGRA in Japan. Several new gravitational wave detectors are planned for the future. These include (i) the Cosmic explorer, which will be similar to LIGO but with 40 km long arms, and (ii) the Einstein telescope, which will have three 10 km long arms in a triangular configuration and will be completely underground. In addition, three space borne observatories are also planned. These are (i) LISA of the European Space Agency, which will have three interferometer arms of length 2.5 million km each, (ii) TianQuin of the Chinese Academy of Science, and (iii) DECIGO, planned by Japan. All three are planned to be launched in the 2030s.

8

Cosmology

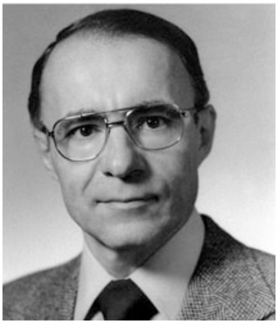
8.1 Introduction

The Nobel Prizes in the category of cosmology were given in the years 1978, 2006, 2011 and 2019 to eight astronomers: Arno Allan Penzias, Robert Woodrow Wilson, John Cromwell Mather, George Fitzgerald Smoot, Saul Perlmutter, Brian Paul Schmidt, Adam Guy Riess, and James Peebles. The award winning work for the year 1978 by Penzias and Wilson was purely experimental. They constructed an antenna in 1964 that immediately resulted in the discovery of cosmic microwave background radiation (CMBR). The discovery paper was published in 1965. The award winning work for the year 2006 by Smoot and Mather was also experimental, resulting in the measurement of the black body shape of the CMBR and the anisotropies therein. It started with a proposal to NASA (National Aeronautics and Space Administration) in 1974. The observations started in 1989. The results were announced in 1990 and 1992 by Mather and Smoot, respectively. The award winning work for the year 2011 by Perlmutter, Schmidt, and Riess was again purely experimental. The work of Perlmutter started with a proposal in 1988. Observations started in 1992 and results were announced in 1997. The work of Schmidt and Riess started in 1994 and results were announced in 1997. Both teams measured distances to distant supernovae, which resulted in the discovery of the accelerated expansion of the Universe. The award winning work for the year 2019 by Peebles is purely theoretical and has been carried out throughout his career spanning several decades. It resulted in an improved understanding of the evolutionary history of the Universe.

P. Khare, *Nobel Prizes in Astronomy*, Astronomers' Universe,
https://doi.org/10.1007/978-3-031-29639-0_8

8.2 Nobel Prize 1978: A. Penzias and R. W. Wilson

The Nobel Prize for the year 1978 was given to Arno Penzias and Robert Woodrow Wilson "for their discovery of cosmic microwave background radiation." They shared half of the prize amount; the other half was awarded to Pyotr Leonidovich Kapitsa that year.



Credit: Penzias: the Leo Baeck Institute

Arno Penzias was born in 1933 in Germany. In 1940 the family migrated to the US. He graduated in physics from the City College of New York in 1954 and worked for two years as a Radar officer in the US army. He joined Columbia University in 1956 and obtained his Ph.D. degree in 1962. After that he joined Bell Labs in New Jersey. He currently serves as a venture partner at New Enterprise Associates.

Credit: Wilson: International Astronomical Union/Lee Pullen

Robert Woodrow Wilson was born in 1936 in the USA. He graduated from Rice University, Houston in 1957 and obtained his Ph.D. from Caltech in

1962. He joined Bell labs in 1963 and was a member of the technical staff. In 1976 he became the Head of the Radio Physics Research Department. In 1994 he shifted to the Harvard–Smithsonian Center for Astrophysics.

8.2.1 Background

Attempts towards theoretically understanding the structure of the Universe were made using Newton's gravity long before Einstein gave his general theory of relativity. Studies based on Einstein's theory got a real boost after Edwin Hubble made interesting observations of several nearby galaxies in 1928–1929. Hubble measured the radial velocities of these galaxies through Doppler shifts of the absorption lines in their spectra, and using Cepheid variables as standard candles (see Box 8.1) he also measured the distances to these galaxies. The radial velocities were expected to be distributed randomly. However, surprisingly, most of the galaxies were found to be moving away from the Milky Way and their velocities of recession from the Milky Way were proportional to their distance from it, as shown in Fig. 8.1. This means that, farther a galaxy is from us, faster is it moving away from us. This relation is known as Hubble's law, and the constant of proportionality between the recession velocities of galaxies and their distances is called Hubble's constant. Does Hubble's law tell us that the Milky Way is special in some sense and is at the centre of the Universe? This seems highly unlikely, in view of the fact that the Milky Way is an ordinary galaxy and there is no reason to believe that it is at the centre of the Universe. Also, even if the Milky Way is indeed at the centre of the Universe, there is no way we can explain the increasing recession velocities of galaxies with their distances. It seems very likely that Hubble's law would be valid even when observations are made from another galaxy. That is, an observer situated in another galaxy, say A, anywhere in the Universe, would also find other galaxies going away from galaxy A, and their velocities of recession from A would be proportional to their distances from A.

Box 8.1 Standard Candles and Cepheid Variables

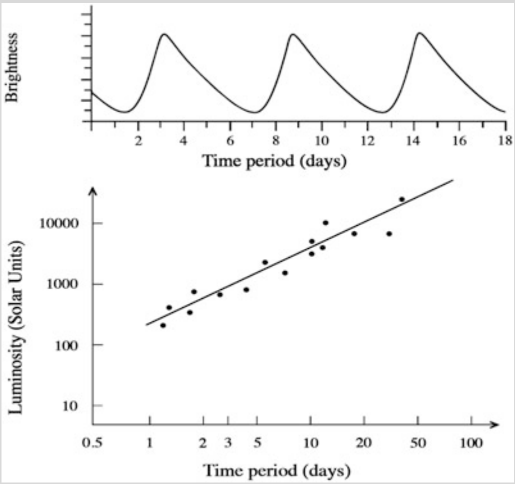
The most popular method for determining distances of distant astronomical sources is the standard candle method. The name comes from the following. If we have a set of standard candles whose luminosity, i.e., the total amount of radiation emitted by a candle per second, has a fixed value, then we can determine the distance of any such candle by measuring its brightness which

(continued)

Box 8.1 (continued)

is the amount of radiation received at the detector, from the source per unit area per unit time. The brightness of a source of given luminosity depends on its distance from the detector. The energy emitted by the source spreads in all directions and the detector only receives a small portion of it. The farther a source is, the smaller is the fraction of its emitted energy that the detector receives per unit area. So if sources having the same luminosity, e.g., standard candles, are located at different distances from us, their brightness will vary, the nearest candle appearing brightest and the farthest candle appearing dimmest. The brightness of a source is inversely proportional to the square of its distance from the detector. This fact can be used to determine distances of standard candles.

What about astronomical sources? If one knows the luminosity of a source and can measure its brightness, then one can determine its distance. Brightness can be measured directly, but how does one determine the luminosity of an astronomical source? With laborious observations of hundreds of astronomical sources and a detailed and tedious study of their properties, astronomers have identified objects whose luminosity can be determined using some of their other observable properties as discussed below. Having obtained the luminosity, the distance can be measured by measuring the brightness. Such objects whose luminosity can be determined from their other observed properties are assumed to be standard candles. What are these objects?



One of the first such "standard candles" to be recognized were the Cepheid variables. These are stars whose brightness varies periodically with time as shown in the upper panel of the figure. The distances of nearby Cepheids, and therefore their luminosities can be measured by using other methods. Studies revealed that there is a direct relationship between the time period of brightness variation and average luminosity, as shown in the lower panel of the figure. Each dot in this figure represents one Cepheid. The straight line is the best fit to the data. This

(continued)

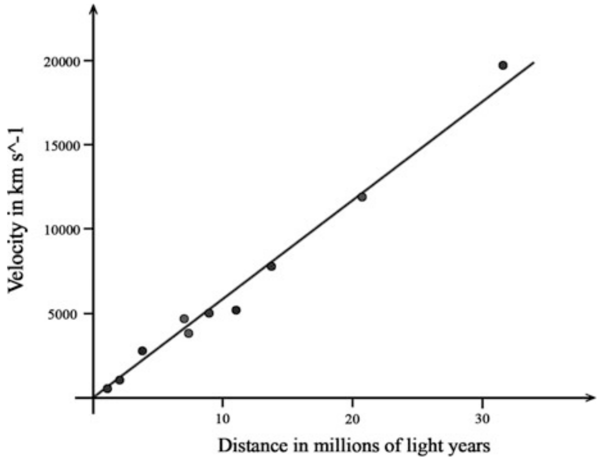


Fig. 8.1 Hubble's law showing a plot of the recession velocities of galaxies versus their distances from the Milky Way. Each dot represents a galaxy and the straight line is the best fit to the data

> **Box 8.1** (continued)
> then makes Cepheids one type of standard candle in the sense that, measuring the period of a Cepheid (by plotting brightness vs. time as in the upper panel of the figure) can yield the average luminosity (using the relation as plotted in the lower panel of the figure). Using the average values of the brightness and luminosity, one can determine its distance. Other examples of astronomical standard candles will be discussed later in this chapter.

Hubble's observations can indeed be explained if the entire Universe, i.e., the three dimensional space making up the Universe, is expanding with time. In this case, the recession velocities of the galaxies are not due to their motion through space, but are because of the fact that space itself is expanding and the galaxies are being carried away from one another due to the expanding space. This can be understood with the example of a balloon described in Box 8.2.

> **Box 8.2 Expanding Space**
>
> To understand what is meant by the expansion of the Universe, we can consider the often used two dimensional analog of the situation which explains the concept beautifully. Think of a balloon without any gas. Assume that the balloon has small equidistant dots printed on it. As we start filling the balloon with gas, the dots start to move away from each other as the rubber between them gets stretched. Consider three dots A, B and C, the distance between A and C being twice that between A and B. As the balloon expands, A and B will drift apart let

(continued)

Box 8.2 (continued)

us say at the rate of 3 mm s^{-1}. The rate at which A and C will drift apart will be 6 mm s^{-1}, as there is twice as much rubber between them which gets stretched as that between A and B. Thus, if we consider from the point of view of A, all dots will be receding away from it and the speeds of recession of the dots (measured along the balloon surface) will be directly proportional to the distances of the dots from it. The situation will be exactly the same when viewed from any other dot on the balloon, as all the dots go away from one another with the expansion of the balloon. One important thing to notice is that the dots are not moving on the surface of the balloon. They are moving apart because the rubber between them is getting stretched.

We can think of an example in three dimensions also. Consider a chocolate chip cake being baked. As the cake gets baked, it expands equally in all directions (assuming there is room for that) and the chocolate chips go away from one another. In ideal case, their motion will also be in conformity with Hubble's discovery for galaxies, i.e., all chips will be going away from one another, and the farther two chips are, the faster will they recede from each other.

Thus, Hubble's discovery can be understood if the three dimensional space around us is expanding uniformly in all directions, so that all galaxies in it are drifting away from one another giving rise to Hubble's law: the velocity of recession of a galaxy, as measured by us, is directly proportional to its distance from us. And this law will be applicable if similar observations are made from any other galaxy as well. It is again worth noting that the galaxies are not moving through space. They are drifting apart because the space between them is expanding. It is also important to understand that the expansion of the Universe is very different from the expansion of the balloon seen above. The two dimensional surface of the balloon expands in the three dimensional space that already exists. The Universe however creates space as it expands as nothing exists beyond the Universe.

The redshift of the spectral lines in the spectrum of a galaxy due to the expansion of the Universe, i.e., due to the cosmic recession velocity, is called the cosmological redshift. The expansion of space stretches electromagnetic waves, increases their wavelengths, and causes redshift of the waves. The motion of galaxies due to the expansion of the Universe is called the Hubble flow. The galaxies may have small motion on top of the Hubble flow due to local effects. This motion is called the proper motion of galaxies. It is important to note that even though space is expanding, the expansion does not affect the structure of gravitationally bound entities like stars, planets, the solar system, a galaxy, or even a cluster of galaxies.

Hubble's law gives us a method for measuring distances to galaxies. One just has to measure the redshift of spectral lines in the spectrum of a galaxy and determine its speed of recession from the Milky Way. Hubble's law then gives us its distance. However, this works only for galaxies which are comparatively close to us. Hubble's linear relation is no longer valid for galaxies at larger distances and the relation gets more complicated. The exact relation can be derived using Einstein's general theory of relativity, but it is model dependent. It depends on the contents of the Universe, as we will see below. The only thing we can be sure is that the higher the redshift of an object, the farther it is from us.

It is obvious that the expanding Universe must have been smaller in the past and the galaxies which are going away from one another must have been closer together (just like the dots on the balloon were closer when there was less air in the balloon and the rubber was less stretched). Assuming the expansion was going on at the same rate in the past, there must have been a time when the material in all the galaxies was present at a single point (similar to the dots on the balloon which were almost at the same point when there was no gas in it). That time would have been the instant when the Universe was born. Hubble's law gives us a way to approximately calculate the age of the Universe at present since its birth. The law gives us the rate at which the galaxies are moving apart. Using it, and assuming, as a first approximation, this rate to have been constant over time, the age of the Universe can, very approximately, be calculated as the distance between any two galaxies divided by their relative velocity. As the relative velocity is directly proportional to the distance, the calculated value of the age of the Universe will be same whichever two galaxies we take. This method gives us the age of the Universe to be close to 14 billion years.

So, working backwards from the present state of the expanding Universe, we know that it had a beginning about 14 billion years back. Obviously the huge amount of matter and energy that is contained in the Universe must have been concentrated at a single point at its birth. The density and even the temperature of the Universe must have been extremely high, tending to infinity. We do not yet have the knowledge of the laws of physics that will be valid under such extreme conditions. However, most scientists do believe that the Universe was born out of a huge explosion, called the "big bang", about 13.8 billion years back as shown by latest observations. The expansion of the Universe caused by this explosion is still continuing even after billions of years that have passed since then.

The Very Early Universe

Our knowledge of physics is inadequate to understand the exact mechanism of the birth of the Universe and its evolution over a short time after that. This is because even Einstein's theory of gravity fails at such high densities, and quantum theory of gravity is needed to understand the behaviour of the Universe. However, we do understand that the density and temperature of the Universe must have started decreasing quickly due to expansion (temperature decrease is similar to that of an expanding ideal gas). The conditions after about 10^{-43} s after the birth no longer remained that extreme and the laws of physics that we have verified in the laboratories are applicable, so we can study the evolution of the Universe with our current knowledge of physics. Even so, the physics is still not completely understood and the evolution of the Universe remains somewhat speculative till about 10^{-11} s after the big bang. Prior to 10^{-11} s, the Universe is believed to have undergone an extremely rapid expansion phase, called inflation, which lasted for a very short time. At these early times, the Universe was filled with elementary particles which are the basic constituents of matter (see Box 2.8). These included quarks and gluons, which are always confined inside particles like the protons and neutrons in the present day Universe. These would have been free at these early times because of the extremely high temperatures and densities. Initial part of the chronology of events after about 10^{-6} s in the life of the Universe, as deciphered by the cosmologists using the well proven laws of physics, is summarized below.

As the temperature of the Universe decreased, at about 10^{-6} s, the quarks and gluons combined to form protons, neutrons etc. The Universe being hot and dense, was filled with a black body spectrum (see Box 2.3) at the temperature of the Universe. This temperature being very high ($\simeq 10^{13}$ K), most of the radiation was in the form of very high energy gamma rays (the peak of the black body curve at such high temperature is in the range of gamma ray wavelengths). With further expansion, the temperature and density decreased and became suitable for nuclear reactions to take place. This period of nuclear reactions, i.e., of big bang nucleosynthesis as we have seen in Chap. 3, lasted from about 10 s to up to 20 minutes. In this period, hydrogen nuclei (i.e., protons) successively combined to form nuclei of light elements like deuterium, helium and lithium. It is worth noting that the most of these light elements, in particular helium, that we see in the Universe, were formed at this time. They were also formed later inside stars but could not come out as such. The observed abundances of these light elements are in good agreement with the predictions of the big bang model and provide one of the major proofs for the model. After 20 minutes in the life of the Universe, the temperature

and density fell to values where nuclear reactions could no longer continue. All the other heavier elements were created much later in stellar nucleosynthesis, as described in Chaps. 2 and 3.

Several assumptions have gone into the big bang theory. The Universe is assumed to have been overall homogeneous, meaning its properties were the same at every point in it, and isotropic, meaning its properties were the same in all directions. In other words, all the points in the Universe are exactly identical. This is of course valid only on large scales, i.e., applies to the average properties of the Universe. On smaller scales we see that matter is not uniformly distributed, but is lumped into stars and galaxies. Only when we take an average of matter density over a large enough volume, is it constant throughout the Universe. Another major assumption is that the laws of physics that we have verified in the laboratory hold good everywhere in the Universe, and at all times (except possibly in the first fraction of a second as noted above). Alternative models to the big bang theory have been proposed, the steady state theory being the most prominent one. According to this theory, the Universe was always the same as it is now and will also remain the same at all time in future. This theory proposes the ongoing creation of matter for the Universe to remain the same in spite of its expansion. The vast majority of astronomers believe in the big bang model of the Universe, as the CMBR and abundances of light elements are consistent with its predictions.

8.2.2 Award Winning Work: Discovery of Cosmic Microwave Background Radiation

In 1964, Penzias and Wilson started working on detecting microwaves which had bounced off ECHO Balloon Satellites. These satellites were passive reflectors of microwaves, and were launched in 1960 and 1964 to help transmit microwave signals from one place on the Earth to another by reflecting them off these satellites. For this purpose, Penzias and Wilson built a 6 m long reflecting horn antenna, which is shown in Fig. 8.2. They made observations at a wavelength of 7.35 cm. As the reflected radio waves that they planned to detect were very feeble, all sources of background noise had to be identified and the background noise had to be eliminated from the received signal. Penzias and Wilson were able to remove the effects of radar and radio broadcasting, and were also able to suppress the interference from the heat in the receiver itself by cooling it with liquid helium to −269°C, a temperature very near absolute zero.

In spite of all this, they found that a steady, low noise continued to appear in their data. This noise was much higher, about 100 times stronger, than the

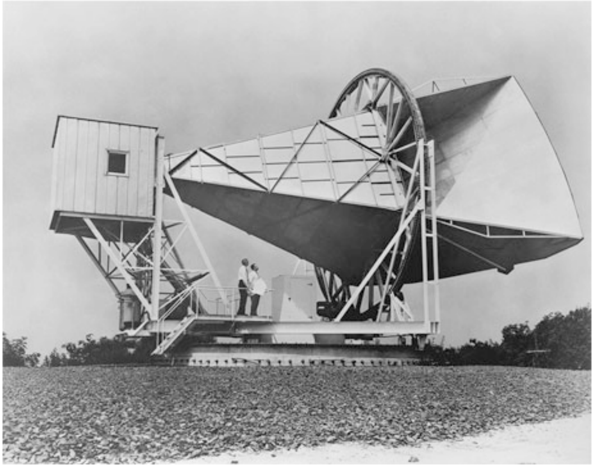


Fig. 8.2 The horn antenna built by Penzias and Wilson. Using this, they discovered cosmic microwave background radiation. **Credit:** By NASA, restored by Bammesk—Original version at Flickr: NASA on The CommonsThis file was derived from: Horn Antenna-in Holmdel, New Jersey.jpeg, Public Domain, https://commons.wikimedia.org/w/index.php?curid=89510549

noise from all known sources that could have been expected. Detailed study showed the noise to be coming from all directions in the sky and was also present at all times, throughout the day as well as night. They concluded that the source of the noise was not on the Earth, the Sun or the Milky Way galaxy. The noise had to be coming from some extragalactic sources. They were not aware of any such sources which could produce the observed noise.

At that time, they heard about the work of Robert H. Dicke, James Peebles, and David Wilkinson, from Princeton University. They were working on the big bang theory. As seen above, according to this theory, during the course of evolution of the Universe, the temperature of the Universe, and therefore that of the background black body radiation, continued to decrease with time and the radiation continuously shifted from being mostly high energy gamma rays to X-rays, ultraviolet rays, and so on. At about $3\text{-}4 \times 10^5$ years after the big bang, the conditions became suitable for the protons and the nuclei (produced during the big bang nucleosynthesis) to combine with electrons and form neutral atoms. Neutral matter cannot interact efficiently with radiation. Thus, the black body radiation which was interacting strongly with matter when the Universe consisted mostly of charged particles, stopped interacting and both radiation and matter continued their further journeys independently. This epoch is called the epoch of decoupling of radiation and matter, both components becoming independent of each other. The radiation continued to move freely. It can be shown that after decoupling, the radiation retained its black body shape but its temperature kept decreasing. In 1941, McKellar through observations and detailed study of cyanogen (CN) and methyne (CH)

molecules in the interstellar clouds, had suggested the presence of a black body background at a temperature of a few degrees above absolute zero, i.e., 0 K.

Peebles, in a paper published in 1965, predicted that the detection of a background black body radiation with temperature as low as 3.5 K, along with the observed abundance of helium can provide the much needed evidence for the big bang model of the Universe. At this temperature, most of the black body radiation (close to the peak of the black body curve) falls in the microwave range and the relic radiation is called the Cosmic Microwave Background Radiation (CMBR). Dicke and Wilkinson had in fact constructed a radiometer, which is an instrument that can measure the intensity of radiation, and a receiving horn antenna capable of detecting this radiation at a wavelength of 3 cm. Penzias and Wilson heard of this and discussed their observations with the Princeton astronomers. They soon realized that they had indeed observed the CMBR. They all decided to publish their results simultaneously in separate papers in the Astrophysical Journal Letters. In their paper, Penzias and Wilson simply reported the observations of the background at a temperature of 3.5 K without speculating on its nature and origin. The accompanying paper by the Princeton astronomers explained this observed radiation to be the cosmic microwave background radiation. The paper by Penzias and Wilson was hardly 2 pages long, but the discovery had enormous importance and earned them the 1978 Nobel Prize.

8.3 Nobel Prize 2006: J. C. Mather and G. F. Smoot

The Nobel Prize for the year 2006 was given to John C. Mather and George F. Smoot "for their discovery of the black body form and anisotropy of the cosmic microwave background radiation."

Credit: Mather: By Christopher Michel–John Mather, CC BY 2.0, https://commons.wikimedia.org/w/index.php?curid=39209520

John Mather was born in 1946 in the US. He graduated in 1968 from Swarthmore College, Pennsylvania and obtained his Ph.D. from the University of California, Berkeley in 1974. He spent two years as an NRC postdoctoral fellow at the Goddard Institute for Space Studies, New York City and then went to the Goddard Space Flight Centre and continued there till 1995. Since then he is a senior Project Scientist for the James Webb Space Telescope.

Credit: Smoot: By Nomo michael hoefner http://www.zwo5.de—Own work, CC BY 3.0, https://commons.wikimedia.org/w/index.php?curid=7131137

George F. Smoot was born in 1945 in Florida, USA. He obtained dual Bachelor's degrees in mathematics and physics in 1966 from the Massachusetts Institute of Technology and continued there to obtain his Ph.D. in particle physics in 1970. In 1971 he joined the University of California, Berkeley. He has been an astrophysicist at Lawrence Berkeley National Laboratory since 1974 and professor at the University of California, Berkeley since 1994.

8.3.1 Background

The discovery of the cosmic microwave background radiation with a temperature of 3.5 K by Penzias and Wilson in 1965 opened up a new era of cosmology, transforming it from speculative to real science. As we have seen, the Universe is assumed to be homogeneous and isotropic and so, every point in it is exactly identical to every other point in it. Thus, the microwave background coming to us from any direction in the sky is expected to have the same black body temperature. We see around us that the matter in the Universe is not uniformly distributed throughout its volume. It is lumped together at places in the form of galaxies. How did this happen? It is believed that for galaxies to form, some density inhomogeneities must have been present in the matter in the Universe early on, i.e., there should have been regions where matter density was higher than its average value. These regions would have higher gravity than their surroundings due to their higher than average mass density, and would attract more and more matter and grow in mass, ultimately forming the structures that we see today as galaxies. These inhomogeneities can influence the CMBR on small scales and can cause small changes in its temperature. The differences in the CMBR black body temperatures in radiation coming from different directions are called anisotropies. The main causes for anisotropies are as follows.

1. Anisotropy produced before the decoupling of radiation and matter: Density inhomogeneities present before decoupling of matter and radiation cause inhomogeneities in the temperature of radiation. These inhomogeneities present at the time of decoupling must have been preserved because, after decoupling, the radiation did not interact with matter and flowed freely. Considering the size of the density inhomogeneities at the time of decoupling, cosmologists can estimate the angular scale, i.e., the angle between two directions of observation on the sky over which the temperature of the CMBR can be expected to differ. They can also estimate the magnitude of the temperature difference.
2. Anisotropies generated after the decoupling of matter and radiation: After decoupling, the matter in the Universe was mostly neutral and did not interact effectively with CMBR photons, rendering them free to travel unhindered through the Universe. However, there were pockets in space where the matter was again ionized, passage through which could affect the temperature of the CMBR.

   One such effect is the so called Sunyaeve–Zeldovich effect. The galaxies in the Universe are not distributed uniformly, but are mostly found in

clusters. The medium in between the galaxies in a cluster is gaseous and very hot and hence is ionized, i.e., the atoms have their electrons stripped off. Thus, it is a pool of charged particles. These being at a high temperature are able to scatter, (i.e., collide with and deflect), the CMBR photons, imparting some of their energy to the photons. Thus, the temperature of the CMBR coming to us from the direction of a galactic cluster would be higher than that of the radiation coming from other directions in which no clusters are present.

Another effect on the temperature of the CMBR is because of the gravitational redshift predicted by Einstein's general theory of relativity (see Sect. 6.2.1). According to this theory, whenever the CMBR photons pass close to any massive object lying along their paths, their energy gets affected by its gravitational field. This causes a change in the temperature of CMBR coming to us from that direction.

All these effects are extremely weak, and cause a very tiny difference in the CMBR temperature. Extremely sensitive detectors are needed to be able to measure them.

8.3.2 Award Winning Work: Blackbody Shape and Anisotropies of the Cosmic Microwave Background Radiation

In 1974, NASA invited proposals for astronomical missions for a small or medium sized Explorer spacecraft. The proposal by John Mather for an instrument named Far InfraRed Absolute Spectrometer (FIRAS) for observing the spectrum of the CMBR, by George Smoot for an instrument called Differential Microwave Radiometer (DMR) for measuring the anisotropies in the CMBR, and by M. Hauser for an instrument called Diffuse InfraRed Background Experiment (DIRBE) for mapping the background of infrared radiation emitted by the dust particles in the galaxies, were accepted; and the Cosmic Microwave Background Explorer (COBE) satellite carrying these three instruments was launched in November 1989.

The anisotropies in the CMBR at that point were expected to be very small, about one part in 1000, and the DMR had to be sensitive to that extent. However, one also would have to remove the noise due to the ambient temperature which is about 300 K and because of this, the anisotropies are expected to be about one part in 10^5 of the noisy background. Thus, for definite measurement of the anisotropy significantly above the noise, a DMR having a sensitivity of one part in a million needed to be constructed.

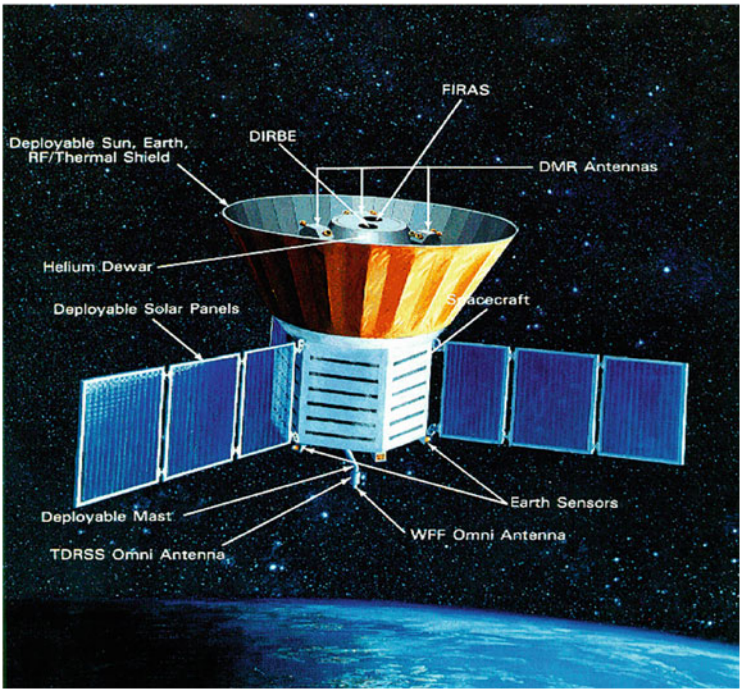


Fig. 8.3 A sketch of the COBE spacecraft. **Credit:** Public Domain, https://commons.wikimedia.org/w/index.php?curid=3888701

Measuring anisotropy meant measuring the difference in the temperature of the CMBR coming from different directions. For this, Smoot decided to use two separate identical antennas which pointed in two different directions, 60 deg apart, in the sky. The instrument could make observations at three wavelengths, namely at 3.3 mm, 5.7 mm, and 9.5 mm. Thus, there were 6 antennas for the three wavelengths, two for each wavelength. The CMBR was expected to contribute strongly at the first two wavelengths, while the signal at 9.5 mm was expected to get maximum contribution from the sources in the Galaxy (Milky Way). This would help in removing the Galactic contribution to the signals received at the other two frequencies. A schematic diagram of the COBE instrument is shown in Fig. 8.3.

As the two detectors would observe the signal coming through part of the Earth's atmosphere, it was necessary that the effect of the atmosphere be same for both antennas and gets cancelled when the signals from the two antennas are compared. As the atmospheric conditions keep changing and very likely would not be identical in the two directions, Smoot decided to

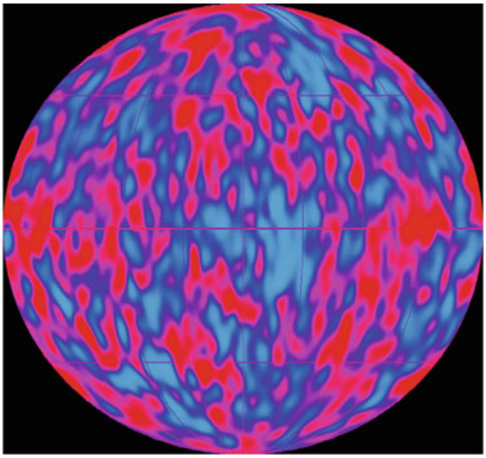


Fig. 8.4 Map of the CMBR temperature as measured by COBE. Different colours indicate different temperatures. The difference in temperatures is of the order of one part in 10^5. **Credit**: The COBE datasets were developed by the NASA Goddard Space Flight Center under the guidance of the COBE Science Working Group. http://lambda.gsfc.nasa.gov/product/cobe/dmr_image.cfmhttp://lambda.gsfc.nasa.gov/product/cobe/cobe_images/cmb_fluctuations_big.gif, Public Domain, https://commons.wikimedia.org/w/index.php?curid=34992487

switch the directions of the two antennas rapidly. Thus, both antennas would observe along both directions in quick succession and the average effect of the atmosphere would be same in both of them. This would ensure that any difference between the signals received from the two directions is indeed due to the anisotropy of the CMBR. In addition, the average background noise had to be removed. All of this required meticulous planning and testing of the instrument using rocket flights. The final instrument aboard the COBE satellite worked very well for 4 years and mapped the whole sky in CMBR. Smoot was able to confirm the presence of anisotropies of the order of 1 part in 10^5, which was consistent with the predictions of the theoretical models. The measured temperatures, indicated by colours, over the entire sky are shown in Fig. 8.4.

FIRAS was constructed by Mather's team, and compared the observed signal with that of an ideal black body installed in the detector itself. It was located at the centre of the COBE assembly, as can be seen in Fig. 8.3. The instrument included an interferometer. It covered wavelength range from 0.1 to 10 mm in two spectral channels. The instrument was cooled to 1.5 K to reduce its thermal emission. The two sources used for interference were the CMBR and the radiation of the reference black body. The optical arrangements were

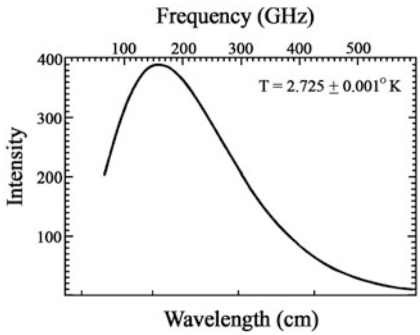


Fig. 8.5 Spectrum of the cosmic microwave background radiation as observed by FIRAS

such that null output would be obtained when the radiations from the two sources matched to better than 1%. By changing the temperature of the reference black body, the CMBR temperature could be obtained precisely. This was the first time the actual shape of the CMBR was obtained with great accuracy. The observations are shown in Fig. 8.5. (Recall that Penzias and Wilson had measured the background radiation only at one wavelength.) The measurements are so accurate that the errors in measurement at any wavelength are smaller than the thickness of the line at that wavelength. The CMBR temperature was found to be 2.735±0.06 K. The deviation from, black body spectrum was shown to be less than 1%. This provided the much needed proof for the big bang model of the Universe. Smoot and Mather were awarded the Nobel Prize for the year 2006 for their remarkably accurate observations of the CMBR.

8.4 Nobel Prize 2011: S. Perlmutter, B. P. Schmidt and A. G. Riess

The Nobel Prize for the year 2011 was given to S. Perlmutter, B. P. Schmidt and A. G. Riess "for the discovery of the accelerating expansion of the Universe through observations of distant supernovae." Half of the prize amount was given to Perlmutter; the other half was shared by Riess and Schmidt.

Credit: The Nobel Foundation, photo: U. Montan

Saul Perlmutter was born in 1959 in the USA. He graduated in 1981 from Harvard University and received his Ph.D. from University of California, Berkeley in 1986. He is an astrophysicist at Lawrence Berkeley National Laboratory and a professor at the University of California, Berkeley.

Brian P. Schmidt was born in the USA in 1967. He graduated from the University of Arizona in 1989 and obtained his Ph.D. from Harvard University in 1993. He was a postdoctoral fellow at the Harvard–Smithsonian Centre for astrophysics in 1993-1994 and moved to Australian National University's Mount Stromlo Observatory in 1995. He became Vice-Chancellor of Australian National University in 2016.

Adam G. Riess was born in the USA in 1969. He graduated from MIT in 1992 and obtained his Ph.D. from Harvard University in 1996. He was a Muller fellow at the University of California, Berkeley and joined the Space Telescope Science Institute in 1999. He later joined Johns Hopkins University in 2006.

8.4.1 Background

As we know, the Universe is expanding and as a result, all galaxies are going away from the Milky Way. The velocity of recession increases with their distance from us, as was first shown by Hubble. This expansion is the result of the big bang that occurred about 13.8 billion years ago, the initial explosion providing the recession velocity to the galaxies. The Universe has matter in it in the form of galaxies and intergalactic matter, and also has energy in the form of radiation. As we know, energy is equivalent to mass, and thus exerts/feels gravitational force. There is a mutual gravitational force of attraction between all the contents of the Universe and so it was expected that this force of attraction would act against the recession of galaxies and their recession velocity must have been decreasing with time after the initial big bang. Very simply, this can be understood by considering the fate of a ball thrown vertically upwards from the Earth's surface, which we have seen earlier. The thrower imparts an initial vertical velocity to the ball. The ball, however, is being continually attracted by the Earth due to its gravitational force. As a result, its velocity keeps decreasing with time and its final fate depends on which one of the two, its kinetic energy (which is taking it away from the Earth but is decreasing continuously with time) or the force of gravity (which is pulling it back towards the Earth and is also decreasing as the distance of the ball above the Earth's surface increases) wins. Thus, the fate of the ball depends

on the initial velocity with which the ball was thrown up, and there are three possibilities.

1. If the initial velocity is larger than the escape velocity (see Sect. 6.2.1), which is about 11 km s^{-1}, the ball will overcome the Earth's gravitational pull and will keep going away from it forever.
2. If the velocity is smaller than the escape velocity of the Earth, the ball will fall back to the Earth after reaching a certain height (where its velocity becomes zero).
3. If its initial velocity is exactly equal to the escape velocity, the ball will just be able to escape the Earth's gravitational attraction, but its velocity will ultimately tend to zero.

The value of escape velocity is decided by the mass and radius of the Earth (which decide the gravitational force felt by an object at the Earth's surface; see Sect. 6.2.1). It would have been smaller/larger than 11 km s^{-1} had the Earth been bigger/smaller or less/more massive than it actually is. Similar to the case of the ball, in case of the Universe, the matter and energy in the Universe try to slow down the expansion of the Universe due to the mutual gravitational attraction, while the initial expansion energy/velocity imparted by the big bang tends to keep the expansion going. Based on this, there appeared to be three options for the future of the Universe.

1. If its initial velocity of expansion was sufficient to overcome the gravitational attraction of all the matter and energy in it, then it will keep expanding forever. This is like case 1 of the ball above.
2. If the initial velocity of expansion was not sufficient to overcome the gravitational attraction of matter and energy in it, then it will stop expanding after a while and start to collapse, reaching the initial state once again. This is like case 2 of the ball above.
3. In case the two forces exactly balance out, the Universe will keep expanding for ever, but the expansion velocity will ultimately go to zero. This is like case 3 of the ball above.

We know the present velocity of expansion of the Universe. So if we can measure the amount of matter and energy that it contains, then we can determine which of the two forces dominates and can predict the fate of the Universe. A new type of matter, which the astronomers named dark matter was discovered in the 1970s (see Box 8.3). Though the nature of dark matter, has not been understood so far, one thing that is definitely known about it is that it exerts and feels gravitational force. Thus, the fate of the Universe depends on the total matter content, including dark matter, and the energy present

in the Universe. In the absence of definite observational knowledge of these quantities, theoreticians construct models of the Universe using Einstein's equations and treating these quantities as parameters. A comparison of the predictions of these models with observed properties of the Universe like the CMBR, abundance of light elements etc., yields the best fit values of these parameters.

It may be noted that whatever be the fate of the ball or the Universe, one definite prediction that can be made in both cases, based on what we have learnt so far, is that the velocity of the ball, as well as the velocity of expansion of the Universe must decrease with time—due to the pulling back by the gravitational force of the Earth in case of the ball, and of the matter, whether dark or visible, and energy, in case of the Universe.

Box 8.3 Dark Matter

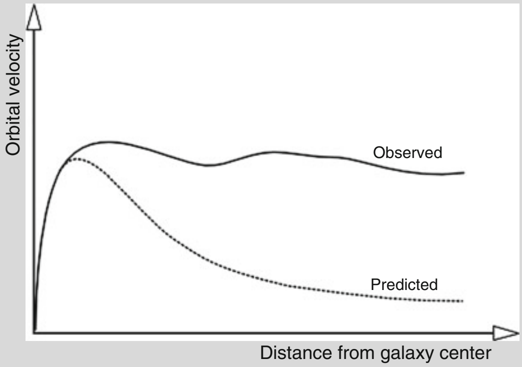
Matter by definition is something that has mass and occupies space. Another important property of matter is that it exerts and experiences force of gravitation. The matter that we see around us like that on the Earth, in the solar system, and in stars, galaxies etc., has another important property as well: it experiences and exerts electromagnetic force. It is because of this property that it is able to emit, absorb, reflect, and scatter electromagnetic radiation and we are able to "see" it. Here, by "see" we mean detect it using any band of the electromagnetic radiation. This matter is, thus, referred to as visible matter.

But there can be another type of matter in the Universe, which does satisfy the first two properties listed above but not the third. That is, while it does have mass, occupies space, and experiences gravitational force, it does not experience the electromagnetic force. Does this type of matter exist in the Universe? We have already come across one such example: the neutrinos. But there is ample, albeit indirect evidence to suggest that additional such matter does indeed exist, and that too in huge quantities. Because such matter does not interact with electromagnetic waves, it can not emit, absorb, scatter, or reflect light or any other form of electromagnetic radiation. As a result, we can never actually "see" this type of matter. How then do we know that it is actually present? It is through the gravitational force that it exerts on its surroundings.

One strong evidence for its existence comes from the so called rotation curves of spiral galaxies. The stars and gas clouds in these galaxies rotate around the centre of the galaxy. The speed of rotation is not constant as that in a rigid body rotation, but depends on the distance of the star or the cloud from the centre of the galaxy. Essentially it depends on the total mass of the galaxy which is contained inside the orbit of the star or the cloud. One can roughly estimate the mass of a galaxy inside a given radius from the amount of light that the galaxy is emitting from that region, assuming that the ratio of a given mass to the amount of light emitted by it is constant and is essentially the same as the mass to light ratio of a star like the Sun. This assumes that all the light of a galaxy is emitted

(continued)

Box 8.3 (continued)
by the stars and that all the mass of a galaxy is in the form of stars. Based on this assumption and on the observed distribution of light emitted by the galaxy as a function of distance from its centre, the expected variation of velocity of rotation of the stars as a function of distance from the centre of the galaxy is shown in the figure. This is very different from the observed velocity distribution which is also shown in the figure. The same behaviour was seen in all spiral galaxies that were observed. This difference can only be explained if the galaxies have much more mass than is indicated by the amount of light emitted by them and that this mass is spread up to larger distances from the centre as compared to the distance up to which the stars are distributed. This additional matter must not be emitting any radiation. This was one major proof for the presence of dark matter, i.e., matter which is not emitting any light.



Further proof was obtained from observations of clusters of stars, from gravitational lensing, etc. No proof of any interaction of dark matter with electromagnetic radiation has been obtained. Most scientists believe that dark matter is in the form of some kind of elementary particles. Many experiments have been designed and conducted to detect such particles through any other type of interactions that they may possess, but the search has not yielded any results so far. This is one of the few big mysteries that is faced by the astronomers today. Comparison of various predictions of theoretical models with observations indicates that the amount of dark matter in the Universe is about 5 to 6 times that of visible matter.

One way to understand the future of the Universe is to determine its past evolution. This can be understood as follows. We know that light takes finite time to reach us from any source and we actually see the source as it was that much time in the past. The farther a source is, the farther back in time do we see it. Using this fact, we can determine the expansion history and the rate of deceleration of the Universe in the past by studying distant

objects. By comparing this with the prediction of the models of the Universe, we can determine the values of the parameters that describe the Universe. Thus, we need to measure the relationship between the distances of bright sources and their cosmological redshifts. The distance of a source would give us the time taken by its light to reach us, and its redshift will tell us how much has the wavelength of light from the source been stretched, i.e., how much the Universe has expanded during that time. By measuring the distances and redshifts of a number of sources distributed over a large volume, i.e., by measuring the distance-redshift relation of sources in the Universe, we can determine the history of expansion of the Universe. This is exactly what Hubble had done. However, he made measurements over very small distances and that gave us just the present expansion rate of the Universe. His observations, being limited to small distances, were not able to determine the expected deceleration of the Universe. Observations of much farther sources would be needed for this purpose. The larger the redshifts and distances of the observed objects, the larger will be the span of time over which we can obtain the expansion history of the Universe and the better equipped we will be to determine the deceleration rate and therefore the future of the Universe. As we know, in order to determine the distances to the sources, we need to know their luminosity. Measuring their brightness, we can then determine their distances. In short, the sources have to be standard candles and have to be bright so that they can be observed over long distances.

Hubble used Cepheid variables to measure distances to nearby galaxies. Cepheids are standard candles (see Box 8.1); however, as they are stars, we can only observe those which are present in sufficiently nearby galaxies, i.e., galaxies which are closer than 2×10^8 light years.

One type of robust standard candle discovered in the 1980s is supernova Type Ia (see Box 8.4). These types of supernovae are extremely bright sources, almost as bright as an entire galaxy when they are at their peak, so that they can be observed up to much larger distances as compared to the Cepheids. Being explosions, their luminosity increases, reaches a peak in about two weeks and slowly decreases over a period of a couple of months. The remarkable thing about these supernovae is that the rate of change in their brightness is related to their peak luminosity. Thus, we can determine their peak luminosity by measuring the rate of change of their brightness over a period of time. Their distances can then be determined from their peak luminosity and their observed brightness. These supernovae are the most reliable of standard candles, meaning that the resulting distances are the most accurate among those obtained using other standard candles considered by astronomers. Also, as they are very bright, they can be seen up to large distances of up to about 10 billion light years.

Box 8.4 Supernova Type Ia

In general, supernovae are a heterogeneous group of exploding stars and their peak intensity varies considerably. Observationally, supernovae are divided in two types. Type I supernovae are those whose spectra do not show any lines of hydrogen. Of these, Type Ia are those whose spectra have a line at 6150 Å produced by silicon. The spectra of all supernovae of this subcategory are very similar and so are their light curves, which are graphs between brightness and time. Such uniformity in the spectra of Type Ia supernovae indicates that the mechanism for triggering their explosions must be the same. The current understanding of the mechanism is as follows.

We have seen earlier (see Box 4.4), how binary systems emit X-rays. There is another spectacular phenomenon involving these systems, which also arises due to the interaction between the two members of the binary. Consider a binary with two stars A and B going round their common centre of mass. Let us assume that one of the stars (say A) is a white dwarf and the other (say B) is a normal star burning nuclear fuel in its centre. During its evolution, when B expands, its outer material on the side of A experiences a larger force due to the gravity of A as compared to that due to the gravity of its own star (B). It therefore gets pulled out and spirals towards A and forms an accretion disc around it. Material from this disc slowly trickles on to A. This material is the source of X-rays in the X-ray binaries, as we have seen in Chap. 4. This material falling on A, results in an increase in its mass. Due to this, a time comes when the mass of A, which is a white dwarf, exceeds the Chandrasekhar limit of 1.44 $M_{\odot}$. The star then cannot maintain its hydrostatic equilibrium as a white dwarf, as the pressure of degenerate electrons can not balance the increased gravity. The star then collapses, causing a continuous increase in temperature and density at the centre As a result, nuclear reactions are triggered in the centre and there is a sudden release of huge amount of energy, leading to a massive explosion which is called Type I supernova explosion. As all such type of explosions originate in stars having masses close to but greater than 1.44 $M_{\odot}$, their light curves and peak luminosities are very similar and render them excellent standard candles. Remember that the Type II supernovae (i.e., the core-collapse supernovae) (see Sect. 3.3.1) occur in stars with initial masses larger than 8 $M_{\odot}$. As these exploding stars have a wide range of masses, their peak luminosities and light curves are very different and they can not be used as standard candles.

However, there are problems in using Type Ia supernovae. Firstly, they are very rare and a typical galaxy may only have a couple of such explosions per millennium. Secondly, one has no clue where and when they will occur. After the explosion, they have to be observed within a few days so as to be able to determine the peak in their light curve. Even if they are discovered early, getting time on the big telescopes which are necessary for their observation (as they are faint, being far away), may not be easy: the observing times on these telescopes are allotted to deserving proposals, months in advance and it is very difficult to get access to big telescopes quickly.

8.4.2 Award Winning Work: Discovery of the Accelerating Expansion of the Universe

It was realized in the nineteen eighties that observations of Type Ia supernovae at large distances can give us important information about the evolutionary history, and therefore the properties, of the Universe. However, as mentioned above, there are problems in observing them. To overcome these problems Perlmutter thought of a strategy. In the late eighties, he along with Carl Pennypacker built a wide-area CCD camera, which could image large areas of sky. This instrument could observe thousands of galaxies in a single night. Comparing the images obtained on two nights, supernova explosions could be detected early. Using this technique, they found the first supernova in 1992. Still the problem of how to guess the occurrence of the Type Ia supernova sufficiently in advance to apply in time for observing run at big telescopes remained. For this, they came up with a plan. They observed several neighbouring patches of sky using their camera just after a new moon. Observing around a new moon night guarantees dark nights and hence enables observation of faint sources. The same patches were observed again just before the next new moon night. Any new or brightened sources indicate supernovae explosions which have taken place after the first set of observations were taken. Simultaneously, they would apply early enough for spectroscopic observing time on large telescopes, the KECK and the VLT, around the second new Moon night, and would by then, have a set of sources confirmed as Type Ia supernovae ready to be observed. They would also apply in advance, for observing time on the Hubble and some other big telescopes to obtain the light curves of these supernovae over the next several months. They demonstrated the validity of this methodology, which they called "Supernovae on demand", in 1994. They named their project as the "Supernovae Cosmology Project". Once they demonstrated the validity of their scheme, they were awarded observing time on big telescopes more easily.

In 1994, another project called the "High redshift supernovae search", essentially having a similar observing strategy, started under the leadership of Brian Schmidt who was joined by Adam Riess.

By 1998, both teams had a sufficient number of sources to be able to understand the evolutionary history of the Universe. Their results were extremely surprising. They could convincingly show that the Universe is

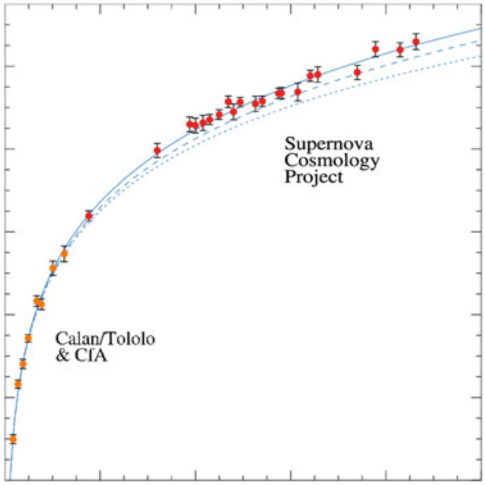


Fig. 8.6 Figure showing the luminosity along the vertical axis and redshift, which is a measure of the distance, along the horizontal axis. The three curves from top to bottom are theoretical predictions for different values of matter content of the Universe as explained in the text. The top curve fits the data best. **Credit:** Knop et al. Astrophys.J. 598, 102, 2003

not decelerating as expected, but is actually accelerating. This was a totally unexpected result. It showed that the density of matter and known type of energy is not only insufficient to slow down the expansion of the Universe but there is some other source of energy which is acting against the gravitational force of matter and radiation and is actually pushing galaxies apart so that they are moving away faster than what was possible due to the initial thrust of the big bang explosion. No such source of energy is known to physicists. We now call this mysterious energy the dark energy. This is something that acts like antigravity and is responsible for the acceleration of the expanding Universe. Figure 8.6 shows a plot essentially between luminosities and distances of a sample of supernovae observed by the Supernova Cosmology Project. The observed data points are shown in red along with their error bars. The blue curves show predictions of theoretical models for three different sets of parameters. The solid curve is the best fit curve and assumes that 25% of the contents of the Universe are in the form of matter and dark matter, while 75% of the contents are in the form of dark energy. The observations thus point to a dark energy dominated Universe which causes its acceleration. The nature of, and the physics behind dark energy is not yet known. The dominance of dark energy indicates that the acceleration of the Universe will continue in the future for all times. Perlmutter, Schmidt, and Riess were awarded the Nobel Prize for the year 2011 for their remarkable discovery.

8.5 Nobel Prize 2019: J. Peebles

The Nobel Prize for the year 2019 was given to James Peebles "for theoretical discoveries in physical cosmology." He received half the prize amount; the other half was shared by M. Mayor and D. Queloz.

Credit: The Nobel Foundation, photo: U. Montan

James Peebles was born in Canada in 1935. He graduated from the University of Manitoba in 1958 and obtained his Ph.D. from Princeton University in 1962, after which he joined as a faculty in Princeton University. He became a full professor in 1972, the Albert Einstein professor of science in 1984, and professor emeritus in 2000. He also spent several years at the Institute of Advanced Study in Princeton. His books, "Physical Cosmology" written in 1971, "The Large-Scale Structure of the Universe" written in 1980, and "Principles of Physical Cosmology" written in 1993 have been widely read and referred to by astronomers.

8.5.1 Background

Einstein had given his theory of general relativity in 1915. His famous equations determine the gravitational field produced by any given distribution of matter and energy. While applying this to study the Universe, the inputs to these equations are the density distributions of matter and of energy at every point at a given time in the Universe and the output is the value of the gravitational field (which in Einstein's theory is described by the curvature of space and time) at all space points at that time. At the time when Einstein gave

his theory, the Universe was believed to be static, as all the observed galaxies and stars appeared to be fixed in space. To his dismay, Einstein found that a static solution was not allowed by his equations. This is mainly because of the gravitational force of attraction between all the matter and energy in the Universe. This force brings matter and hence galaxies, closer and closer, making a static Universe impossible. Einstein, therefore, added one more term to his equations which he called the lambda term. This term is also called the vacuum energy term. This artificially introduces an effective repulsive force between material particles to balance the attractive force of gravity, thereby making a static Universe solution possible.

Soon after this, in 1928, Hubble made his discovery which showed that the Universe was not static and was actually expanding (see Sect. 8.2.1). The expansion was due to the initial big bang explosion at the time of its birth. There was, thus, no need for the lambda term any more. Einstein regretted having introduced it earlier and called it his biggest blunder. However, not much was happening on the observational front after that and there was little interest in cosmology among theoreticians. The next big thing to happen on the cosmological front was the discovery of the CMBR by Penzias and Wilson in 1965. After that, the field developed rapidly, especially after the launch of the COBE satellite and the definitive discovery of anisotropies in CMBR by Smoot.

Cosmology soon became a precision science. Various parameters describing the Universe could be observationally determined to the accuracy of several decimal places and predictions from theoretical models could be verified observationally. Two more satellites, Wilkinson Microwave Anisotropy Probe (WMAP) launched by NASA in 2001 and PLANCK launched in 2009 by ESO further measured the details of the CMBR and its anisotropies. Progress was made on other observational fronts as well with large automated galaxy surveys being undertaken. Samples of high redshift supernovae grew in size.

8.5.2 Award Winning Work: Theoretical Discoveries in Physical Cosmology

Throughout his scientific career, Peebles worked on theoretical cosmology. He laid the foundations for almost all modern investigations in cosmology, both theoretical and observational. His important contributions are summarized below.

In 1965 he calculated the temperature of the background radiation at the present time, and showed that it would peak in the microwave region. He is therefore credited for the prediction of the Cosmic Microwave Background Radiation.

Around the same time, he showed that the black body radiation filling the early Universe exerts radiation pressure on matter and pushes material particles apart. This meant that the density inhomogeneities in the Universe could not grow and galaxies could not form. He showed that only after the Universe had expanded, and as a result, cooled enough so that the radiation pressure decreased sufficiently, would the excess gravity in the density inhomogeneities overcome it. Only then would they grow, making galaxy formation possible.

In 1970, Peebles worked on possible anisotropies in CMBR spectrum. He particularly studied their dependence on the matter density of the Universe. He calculated the growth of anisotropies and predicted their shape at the present time. Anisotropies had not been observed at that time. They were indeed observed, albeit much later, by COBE, WMAP, and Planck satellites, lending support to Peebles' calculations. Peebles was the leading cosmologist in "structure formation, i.e., formation of galaxies in the Universe" during that period.

As we have seen, in the very early stage of the Universe, the temperature and density were extremely high and nuclear reactions could take place. It was not clear which elements would be produced in this nucleosynthesis phase. Peebles showed that the temperature of the Universe had a great effect on the amount of helium produced. He also showed that at an early stage, the temperature of the Universe would have dropped low enough that deuterium would no longer be converted into helium, and thus the big bang nucleosynthesis would stop. He made major contributions to the theory of the big bang nucleosynthesis.

In the 1970s, it was discovered that the rotation curves of spiral galaxies did not match the curves expected on the basis of the observed distribution of the stars, i.e., of mass in these galaxies (see Box 8.3). It led astronomers to suggest the presence of dark matter. However a number of astronomers did not agree with the presence of hypothetical dark matter and suggested that the existing theory of gravity may not be correct and needed modification. Peebles supported the hypothesis of the presence of dark matter which is now believed by most astronomers.

After dark matter was proposed, there were two versions which were considered: hot dark matter and the cold dark matter. Hot dark matter was in the form of particles which would be relativistic in the very early Universe, while cold dark matter is in the form of particles with much lower speeds. Peebles was one of the first cosmologists to show that the presence of cold

dark matter is crucial to the formation of structures such as galaxies and galaxy clusters, ruling out hot dark matter.

He was awarded the Nobel Prize in 2019 for his enormous contribution to all aspects of theoretical cosmology.

Appendix

List of Boxes

P. Khare, *Nobel Prizes in Astronomy*, Astronomers' Universe,
https://doi.org/10.1007/978-3-031-29639-0

Glossary

Absorption lines Dark lines in the spectrum of a star or a galaxy, produced by absorption of radiation traveling from the interior of the object to the outside by matter present along its path.

Accretion disc A disc formed by diffuse material in orbital motion around a star.

Advanced LIGO/aLIGO An upgraded version of the initial LIGO detector. Its sensitivity is about 10 times that of the initial LIGO.

Alpha particle The nucleus of helium, having two protons and two neutrons.

Angular diameter The angle subtended by two diametrically opposite points of an astronomical object at the observer's eye.

Anisotropies Differences in the temperature of CMBR coming from different directions.

Aperture synthesis A type of interferometry, that mixes signals from a collection of telescopes to produce images having the same angular resolution as a telescope having a much bigger aperture.

Atomic number The number of protons in the nucleus of an atom of an element.

Big bang nucleosynthesis The generation of light elements in the first few minutes in the life of the Universe.

Big bang theory Theory describing the origin of the Universe in an explosion and its evolution from an initial state of high density and temperature.

Binary pulsar A pulsar in a binary system.

Binary system (Binary stars) A gravitationally bound pair of stars orbiting around each other.

Black body An object that absorbs all electromagnetic radiation falling on it. The nature of the radiation emitted by such a body depends only on its temperature.

Black hole An object whose gravity is so strong that not even electromagnetic radiation can escape from it.

P. Khare, *Nobel Prizes in Astronomy*, Astronomers' Universe,
https://doi.org/10.1007/978-3-031-29639-0

Blueshift Decrease in the wavelength of electromagnetic radiation during its passage from a cosmic object to an observer. The decrease can occur due to the relative motion of the emitter towards the observer, which is known as the Doppler effect, or can be due to gravitational effects.

Brightness The amount of light received from an astronomical source per unit area per unit time.

Cepheid variable A type of star whose luminosity varies periodically with a characteristic pattern, due to the periodic oscillations of the surface of the star.

Chandrasekhar limit The maximum mass that a white dwarf can have.

Chemical energy The energy of chemical substances that is released when they undergo a chemical reaction and transform into other substances.

Chromosphere The second outermost layer of the Sun, located between the photosphere and the corona. It is several thousand kilometres thick.

CMBR Cosmic Microwave Background Radiation.

CNO cycle A chain of nuclear reactions converting hydrogen to helium. This chain involves nuclei of carbon, nitrogen, and oxygen and takes place at the centres of stars having masses above about 1.3 solar masses.

COBE Cosmic Microwave Background Explorer, a satellite launched in 1989 to determine the spectrum and anisotropies in the CMBR.

Compact objects Cosmic objects which are stellar remnants produced at the end stages of a star's life. They have a radius much smaller than the radius of a normal star, and have very high density. These include white dwarfs, neutron stars, and black holes.

Core-collapse supernova Spectacular explosions of stars with masses larger than about 8 solar masses, at the end of their lives, giving birth to neutron stars and black holes.

Corona The outermost layer of the Sun, reaching temperatures of about a million Kelvin and extending to a few solar radii.

Cosmic Explorer A laser interferometric gravitational wave detector which is being planned. It will have 40 km long arms and will be located in the USA.

Cosmic Microwave Background Radiation Electromagnetic radiation which is a remnant from an early stage of the Universe; also known as "relic radiation", and filling up entire space.

Cosmic rays High-energy protons and atomic nuclei that move through space at nearly the speed of light. They originate from cosmic sources.

Cosmological redshift Redshift caused by the expansion of the Universe.

Dark energy A type of energy present in the Universe, which acts opposite to gravity and exerts a repulsive force which causes matter to go apart, resulting in the accelerated expansion of the Universe.

Dark matter A type of matter which interacts through gravitational force but not through electromagnetic force. Such matter can not emit, absorb, reflect, or scatter electromagnetic waves, and so can not be "seen". Its presence can only be detected through its gravitational influence on the surroundings.

DECIGO The DECi-hertz Interferometer Gravitational-wave Observatory. It is a planned Japanese space-borne mission to detect gravitational waves in the frequency range of 0.110 Hz.

Degenerate gas High density electron or neutron gas, for which quantum mechanical effects are important.

Degeneracy pressure Pressure due to high density electron or neutron gas caused by quantum mechanical effects.

Density inhomogeneities Regions in the early Universe having matter density higher than its average value. These grew to form the structures like galaxies that we see today.

DIRBE Diffuse InfraRed Background Experiment, on board the COBE satellite.

DMR Differential Microwave Radiator, on board the COBE satellite.

Doppler effect The change in wavelength of a wave, as measured by an observer who is having nonzero radial velocity relative to the source.

Einstein telescope A laser interferometric gravitational wave detector being planned, which will have three 10 km long arms in a triangular configuration and will be located in Europe.

Einstein's equations Equations derived by Albert Einstein, which determine the gravitational field for given matter and energy distribution.

Electromagnetic waves Waves having periodically varying electric and magnetic fields.

Elementary particles Basic constituents of nature from which all matter is built.

Escape velocity The minimum velocity with which a mass has to be thrown outward from the surface of a cosmic object for it to escape the gravitational field of the object.

Event horizon An imaginary closed surface surrounding a black hole, from inside of which no signal can emerge to the outside world.

Extrasolar planets (exoplanets) Planets revolving around stars other than the Sun.

FIRAS Far InfraRed Absolute Spectrometer, on board the COBE satellite.

Fission A nuclear reaction in which the nucleus of an atom splits into two or more smaller nuclei.

Free particle A particle not influenced by any force.

Fundamental particles Elementary particles.

Fusion A nuclear reaction in which two or more atomic nuclei combine to form one or more different atomic nuclei and subatomic particles.

Galaxy A gravitationally bound system of stars, stellar remnants, interstellar gas, dust, and dark matter.

Gamma rays The most energetic part of electromagnetic spectrum, having wavelength smaller than 0.01 nm.

Gamma-ray burst Cosmic bursts of highly energetic gamma rays, lasting from less than a second to several minutes.

General theory of relativity Theory of gravity given by Einstein in 1915, in which gravity is manifested as curvature of space-time.

Gravitational lensing Distortion and magnification of light, caused by the bending of rays of light by a gravitational field. This was predicted by Einstein's general theory of relativity.

Gravitational potential energy The stored energy due to the gravitational interaction between two masses. Conventionally this is taken to be negative. The greater the gravitational force between two masses, the more negative is the energy between them.

Gravitational redshift Increase in wavelength of electromagnetic radiation while moving from a region of a strong gravitational field to a region of a weak field.

Gravitational time dilation The fact that a clock placed in a strong gravitational field runs slower than an identical clock placed in a weaker gravitational field. This is a prediction of the general theory of relativity.

Gravitational waves Periodic disturbances in the curvature of space-time, i.e. in the gravitational field, generated by accelerated masses, that propagate as waves outward from their source at the speed of light.

Homestake experiment An experiment set up at Homestake mines by Raymond Davis, to detect neutrinos emitted in nuclear reactions taking place at the centre of the Sun.

Hubble's law Law discovered by Hubble, according to which the speeds of recession of galaxies with respect to us are directly proportional to their distances from us.

Hydrogen burning Conversion of four hydrogen nuclei, i.e., protons into a helium nucleus, i.e., alpha particle, taking place inside a star.

Hydrostatic equilibrium The state of equilibrium of a fluid between external forces like gravity and the force due to pressure gradient.

Ideal gas A gas whose particles occupy negligible space and do not interact with one another.

Infrared radiation Electromagnetic radiation with wavelengths in the range of about 700 nm to about 1 mm.

Initial LIGO/iLIGO The LIGO detector which was functioning till 2015. This has been superseded by Advanced LIGO.

Interference Phenomenon in which two waves superpose to form a resultant wave of greater, lower, or the same amplitude.

Interference pattern The pattern in intensity produced due to interference.

Interferometry A technique which uses the interference of two waves to extract information.

Intergalactic matter Matter existing in the space between galaxies.

Interstellar matter Matter existing in the space between stars.

Ionosphere The ionized part of the upper atmosphere of the Earth, from about 48 to 965 km above sea level.

Isotopes Nuclei with the same numbers of protons but different number of neutrons.

KAGRA The Kamioka Gravitational Wave Detector, a laser interferometer located underground, in the Kamioka mine in Japan.

KAMIOKANDE The Kamioka Nucleon Decay Experiment in Japan, set up by Masatoshi Koshiba to detect neutrinos.

Kepler's laws Three laws given by Johannes Kepler in the seventeenth century, describing the motion of planets around the Sun.

Kerr black hole A rotating black hole. The solution of Einstein's equations which describes such a black hole was discovered by Roy Kerr.

Large Magellanic Cloud A satellite galaxy of the Milky Way, at a distance of about 160,000 light years

Laser A device which emits light through a process known as the stimulated emission of radiation. The emitted radiation is coherent, has a specific wavelength, and spreads very little as it travels forward. A laser beam is emitted by a laser device.

Light curve A graph between the brightness of a variable source and time.

LIGO Laser Interferometric Gravitational Wave Observatory. The Observatory consists of a LIGO gravitational wave detector located in Hanford, Washington State and another at Livingston, Louisiana.

LIGO Scientific Collaboration A collaboration of scientists who carry out scientific and technical work related to LIGO.

LISA The Laser Interferometric Space Antenna, a future space-borne gravitational wave observatory.

Luminosity The total energy emitted by a source per second.

Mass number The number of nucleons in the nucleus of an atom of an element.

Maxwell–Boltzmann distribution The distribution of speeds of particles of an ideal gas.

Michelson's interferometer A precision instrument that produces interference fringes by splitting a light beam into two parts and then recombining them after they have traveled different optical paths. It is used for accurately measuring distances or changes therein.

Neutron star A compact object made mostly of neutrons, in which neutron degeneracy pressure balances gravity. It is the end stage of intermediate mass stars.

Newton's laws of motion The laws that describe the relations between velocity, acceleration, and the distance traveled by a particle.

Nuclear binding energy The minimum amount of energy required to disassemble a nucleus into its constituent particles,

Nuclear energy The energy released when nuclei undergo a nuclear reaction like fission or fusion, because the total mass of the reactants is larger than the total mass of the products.

Nucleon Protons or neutrons present in a nucleus.

Perihelion The point along the orbit of a planet, asteroid, or comet that is nearest to the star at the focus of the orbit.

Photon The smallest discrete amount of energy of electromagnetic radiation. It is the particle form of electromagnetic waves. The energy of a photon is proportional to the frequency of the radiation.

Photosphere The surface layer of the Sun, the light emitted from which can come out of the Sun mostly unabsorbed.

Piezoelectric crystal A material which generates electric charge when pressure or strain is applied to it.

Planetary nebula A type of emission (light emitting) nebula consisting of an expanding, glowing shell of gas ejected from stars with masses between 1 to 8 solar masses toward the end of their lives.

PP chain A chain of nuclear reactions converting hydrogen to helium. This chain takes place at the centres of stars having masses below about 1.3 solar masses.

Precession A slow change in the orientation of the orbit of one body revolving around another, like the orbit of a planet around the Sun.

Pulsar Highly magnetized rotating neutron star, which emits electromagnetic radiation and particles in two narrow oppositely directed beams along its magnetic axis.

Quantum mechanics The branch of physics applicable to the motion and interactions of atomic and subatomic particles.

Quasar A very compact cosmic object which emits an exceptionally large amount of energy over the entire range of the electromagnetic spectrum. The emission is believed to be due to matter falling on to a supermassive black hole at the centre of a large galaxy.

Quasi-stellar objects Quasars.

Radial velocity The relative velocity between a source and an observer along the line joining them.

Radiometer An instrument to measure the intensity of electromagnetic radiation.

Redshift Increase in wavelength of electromagnetic radiation during its passage from a cosmic object to an observer. The increase can occur due to the relative motion of the emitter away from the observer, which is known as the Doppler effect, or can be due to gravitational effects or cosmic expansion.

Resolving power The ability of a telescope to resolve two nearby sources.

Ringdown The phase through which two merged compact objects settle down to a single black hole.

Rotation curve Graph between the speed of rotation of stars and gas clouds in spiral galaxies, and their distances from the centre of the galaxy.

Schwarzschild radius Distance of the event horizon from the centre of a nonrotating black hole which is associated with the Schwarzschild solution.

Schwarzschild solution Solution of Einstein's equations for a massive, nonrotating point source. It was discovered by Karl Schwarzschild in 1916.

Scintillation Atmospheric effects which cause the twinkling of stars.

Seismic noise Disturbances in the positions of mirrors of gravitational wave detectors, caused by seismic disturbances including those caused by human beings, winds, and tidal motions in the ground caused by the Sun and the Moon.

SETI Search for ExtraTerrestrial Intelligence. It is a nonprofit research organization.

Singularity A point where the gravitational field and the curvature of space-time become infinite.

Special theory of relativity A theory of the relationship between space and time developed by Albert Einstein in 1905.

Spectrum The spread of radiation across a range of wavelengths. For electromagnetic waves, this is known as the electromagnetic spectrum. A spectrum can have a continuous spread of energy over a range of wavelengths, as well as discrete features at specific wavelengths.

Standard candle An astronomical source with known luminosity.

Stellar nucleosynthesis The process involving nuclear reactions in stars through which new atomic nuclei are synthesized from pre-existing nuclei or nucleons.

Strain The ratio of the change in the arm length of the LIGO detector, caused by the passage of a gravitational wave, to the original length of the arm.

Supermassive black holes Black holes with masses from hundreds of thousands to billions of solar masses. These are believed to be present at the centres of most large galaxies.

Supernova explosion The powerful and luminous explosion of a star towards the end of its life cycle, or of a white dwarf in a binary system, resulting in a neutron star or a black hole.

Supernova remnant The structure resulting from the explosion of a star in a supernova.

Thermal energy Internal energy possessed by an object in the form of heat.

Thermal radiation Radiation emitted by a body on account of its thermal energy.

TianQin A space-borne experiment for the detection of gravitational waves in the 0.1100 mHz band, being built by the Chinese Academy of Science.

TOV limit An upper limit on the mass of neutron stars.

Triple alpha reaction The reaction between three alpha particles to produce a carbon nucleus.

Type Ia supernova A type of supernova that occurs in binary systems in which one of the stars is a white dwarf. Matter from the companion star falls on the white dwarf which explodes on reaching a mass larger than the Chandrasekhar limit.

Type II supernova Core collapse supernova.

Ultraviolet radiation Electromagnetic radiation with wavelengths in the range of about 100 to about 400 nm.

VIRGO A laser interferometer for gravitational wave detection located near Pisa in Italy.

Visible radiation Electromagnetic waves to which our eyes are sensitive. This has wavelength of about 380 to about 740 nm.

White dwarf The end stage of low mass stars. The gravity of the white dwarf is balanced by electron degeneracy pressure to provide it a stable structure.

X-rays High energy electromagnetic radiation with wavelength in the range of about 0.001 to about 10 nm.

X-ray binary An X-ray emitting binary system consisting of a compact object and a normal star. The X-rays are emitted when some matter from the normal star is transferred to the compact object, with the matter becoming hot enough to emit X-rays.

Index

P. Khare, *Nobel Prizes in Astronomy*, Astronomers' Universe,
https://doi.org/10.1007/978-3-031-29639-0